

بناء وتصميم الاختبارات والمقاييس النفسية والتربوية

الدكتور
عمر طالب الريماوي
أستاذ علم النفس المشارك



بناء وتصميم الاختبارات والمقاييس النفسية والتربوية

بناء وتصميم الاختبارات والمقاييس النفسية والتربوية

كتاب محكم علمياً

الدكتور

عمر طالب الريماوي

أستاذ علم النفس المشارك

الطبعة الأولى

2017م



دار امجد للنشر والتوزيع

المملكة الأردنية الهاشمية
رقم الإيداع لدى دائرة المكتبة الوطنية
(2016/6/3698)

378.5657

الريماوي ، عمر طالب

بناء وتصميم الاختبارات والمقاييس النفسية والتربوية/ عمر طالب الريماوي، عمان ، دار أمجد للنشر والتوزيع،2016.

( ) ص

ر.إ: 2016/6/3698

الواصفات:/ الاختبارات// التقويم التربوي// علم النفس التربوي

ردمك : ISBN:978-9957-99-336-4

Copyright ©

جميع الحقوق محفوظة: لا يسمح بإعادة إصدار هذا الكتاب أو أي جزء منه أو تخزينه في نطاق استعادة المعلومات أو نقله بأي شكل من الأشكال، دون إذن خطي مسبق من الناشر.

All rights reserved. NO Part of this book may be reproduced, stored in aretrival system, or transmitted in any form or by any means, without prior permission in writing of the publisher.

دار أمجد للنشر والتوزيع

جوال :
٠٠٩٦٢٧٩٦٩١٤٦٣٢
٠٠٩٦٢٧٩٩٢٩١٧٠٢
٠٠٩٦٢٧٩٦٨٠٣٦٧٠

هاتف :٠٠٩٦٢٦ ٤٦٥٢٢٧٢

فاكس:٠٠٩٦٢٦ ٤٦٥٣٣٧٢

dar.almajd@hotmail.com
dar.amjad2014dp@yahoo.com

عمان - الأردن - وسط البلد- مجمع الفحيص - الطابق الثالث



كلمة شكر، وإهداء

في البداية يسعدنا أن أشكر السيد الأستاذ الدكتور ماهر اسماعيل صبري رئيس رابطة التربويين العرب، على حسن تعاونه والاهتمام والمتابعة معي. كما لا يفوتني أن أشكر السادة الأساتذة المحكمين المحكم أ.د نادية عبده أبو دنيا أستاذ علم النفس التربوي ومدير وحدة القياس والتقويم بكلية التربية - جامعة حلوان. والاستاذ الدكتور أ.د. رمضان محمد رمضان، أستاذ ورئيس مجلس قسم علم النفس التربوي كلية التربية- جامعة بنها الذين بذلوا جهدهم في قراءة الكتاب والإدلاء بملاحظاتهم ... و لكل من شارك في خدمة هذا الكتاب وتحضيره.

اهداء

إلى زوجتي الغالية ربما التي لم تتوانى في توفير سبل الراحة لي، وصبرت وتحملت الصعاب خلال رحلتي لإعداد هذا الكتاب، وإلى أبنائي الأحبة لتحملهم غيابي عنهم لفترات طويلة ...... فلذات أكبادي: أبنائي الأعزاء (ريتا، ونتلي، ورامي) لتحملهم انشغالي عنهم، فهم شمعة حياتي، أسأل اللـه أن لا يحرمني من وجودهم في حياتي .

مقدمة المؤلف:

يسرني أن أقدم لكل من يهتم بموضوع القياس النفسي، وكل مشتغل بالاختبارات والمقاييس . حيث يلعب القياس النفسى دورا بالغ الاهمية وقد عملت جاهداً إلى أن أجمع أطراف هذه الموضوعات من واقع الخبرة والممارسة، ويعد القياس من أدوات البحث العلمي الأكثر شيوعاً في البحوث العلمية حتى أنه أصبح الأداة الأولى التي تجمع بها المعلومات التي يمكن على ضوئها اختبار فروض الدراسة. حيث ان تصميم أداة البحث، يجب أن تتناسب مع طبيعة المشكلة ومنهج البحث المستخدم والعمر الزمني للمفحوصين.

وذلك لأن كل أداة لها خاصيتها تلزم الباحث للأخذ بها دون غيرها، لأن لكل أداة مميزاتها وعيوبها والمسلمات التي تستند اليها، لذا لابد أن يعي الباحث متى يستخدم الأداة المناسبة للموضوع، وهذا الكتاب يضع أمام الباحث مجموعة من المهارات والقواعد الاساسية والمنهجية المفيدة واللازمة لبناء الاختبارات والمقاييس. لذلك فانه ينبغى على الباحث أو الاخصائى النفسى أن يكون على دراية كاملة بكيفية تطبيقها على النحو السليم.

ويضم هذا الكتاب خمسة فصول: اولاً: يتناول الفصل الأول، المفاهيم الأساسية المتصلة بالقياس النفسي والتربوي، بحيث تمكن الدارسين والباحثين والممارسين من التواصل فيما بينهم بلغة مشتركة متفق عليها. نشأة وتطور تاريخ الاختبارات والمقاييس، وتصنيف الاختبارات كأداة للتشخيص النفسي، استخداماتها كأدوات ذات أغراض متعددة، يصعب حصرها نظراً لتنوع

الظـروف والمواقـف، التـي تتطلـب تطبيـق هـذه الأدوات للاسـتناد إلى نتائجهـا في اتخـاذ قرارات. ويتناول الفصل الثاني: مراحل تصميم وإعداد الاختبار، ومـدى مناسـبة كـل مـنهما للأغراض المختلفة، والأهداف التي يوضع من أجلها الاختبار، أسلوب إعداد الاختبار صياغة بنود الاختبار بشكل علمي، العناصر والخطوات الاساسية في تصـميم الاختبـار. أمـا الفصـل الثالث: يتناول اهم الخصائص السـيكومترية للاختبـارات، الصـدق، والثبـات، والموضـوعية وانواعها.الفصل الرابع: يتناول هذا الفصل المعاييـــر، والمتغـيرات : التـي تتضـمن كراسـة تعليمات الاختبار، التي تبين بوضوح الدرجات التي يحصل عليها المختبر في عينات التقنيين، وعملية إستخدام الإحصاء لوصف أداء المجموعات عـلى الاختبـار، التـي تواجـه الأخصـائي النفساني في قياسه، أو في تفسيره لخصائص السـلوك الإنسـاني، وشروط تطبيقهـا وتعلـيمات الإجراء، واحتساب الدرجة وغيرها، كـذلك يتنـاول مسـتويات القيـاس وانواعهـا، كـما يـتم التطرق الى أخلاقيات اسـتخدام الاختبـارات النفسـية. أمـا الفصـل الخـامس: يتنـاول اهـم المفاهيم والمصطلحات المستخدمة في القياس النفسي والتربوي. وأخيرا فإنني أرجو أن يجـد القارئ في هذا الكتاب كل ما يمكن أن يساعده على تفهم مادة القياس النفسي .

د. عمر الريماوي

فهرس المحتويات:

الفصل الأول

مقدمة:

لقد حاول علماء النفس التمييز العلمـي الـدقيق بـين شخصـيات الأفـراد باتبـاعهم الطـرق والاساليب الكمية في القياس، فعالم النفس يود معرفة مقدار سمة شخصية معينة لدى فرد ما، أو السمات التي تميزه عن غيره من الافراد. وربما كان افتقار بعض العلوم في المـاضي إلى مثل هـذه الأسـاليب الموضـوعية في القيـاس هـو الـذي أدى إلى التبـاين الواضـح والجـدل المستمر حول تعريف المفاهيم، والمصطلحات والقواعد المتعلقة بالظواهر التي تهتم هـذه العلوم بدراستها. ولعل علم النفس يعـد مثـالا واضـحا لـذلك، فدقـة وموضـوعية القيـاس تسهم في تحديـد المفـاهيم، وبلـورة التفكـير مـن أجـل الفهـم المسـتنير لطبيعـة الظـواهر المختلفة، ولما للقياس أهمية كبيرة في أي علم من العلوم، فجميـع العلـوم تسـعى لتطـوير أساليب موضوعية دقيقـة لقيـاس الظـواهر المتعلقـة بهـا مـن أجـل فهـم هـذه الظـواهر وتفسيرها، والتنبؤ بالعلاقات القائمة بين متغيراتها، ومحاولة ضبطها والتحكم فيها. فالتقـدم العلمي يعتمد إلى حد كبير على تمثيل الظواهر والأحـداث وصـياغتها بأسـاليب موضـوعية دقيقة، بحيث تمكن الدارسين والباحثين والممارسين من التواصل فيما بيـنهم بلغـة مشـتركة متفق عليها، كما تمكنهم من التقييم الموضوعي للنتائج التي يتوصل إليها العلـماء، وبـدون هذه الأساليب يصبح تعريف المفاهيم والمصطلحات والقواعد المتعلقة بالظواهر المختلفـة خاضعاً للآراء الذاتية، والأحكام الفردية، وبذلك لا تكون هناك أسس متفق عليها في دراسة وبحث هذه الظواهر، وبدون ذلك تصبح الدراسة العلمية لهذه الظواهر خاضعة للتأملات العقلية والخبرات الشخصية.

مفهوم القياس النفسي والتربوي:

تزخر كتب علم النفس بالكثير من التعريفات، ولعل أشهرها تعريف أنستازي بأنه "مقياس موضوعي مقنن لعينة من السلوك"، في حين يعرفه كرونباخ بأنه " طريقة منظمة للمقارنة بين سلوك شخصين أو أكثر، وقد ناقش أبو حطب هذين التعريفين وتوصل إلى تعريف أكثر شمولاً، فهو يرى أن استخدام كرونباخ لعبارة "طريقة منظمة" في تعريفه للاختبار اكثر دقة من كلمة " مقياس " عند أنستازي، لأنه يتضح من تعريفها للاختبار النفسي أنها لم تميز بين مصطلح اختبار ومصطلح مقياس، فعلى الرغم من تداخل معانيهما، إلا أنهما ليسا مترادفين، فمصطلح مقياس أكثر عمومية من مصطلح اختبار فهو يستخدم في العلوم الطبيعية كقياس الحرارة والوزن والطول والضغط ولايقال على هذه المقاييس اختبار، أما الاختبار يستخدم في كل ميادين البحث السيكولوجي بهدف الحصول على أوصاف كمية، كما هو الحال في بحوث الإدراك والإحساس في التجارب السيكوفيزيائية، والتي تستخدم في هذه الأغراض المقاييس الفيزيائية، فتهتم بقياس الكم الفيزيائي، المقابل للكم النفسي.ويستخدم لفظ اختبار ومقياس بصورة تبادلية في العلوم النفسية فنقول اخبار للذكاء أو مقياس للذكاء .

وليس كل الاختبارات مقاييس، فقد نجد أن بعض الاختبارات لا تعطي درجة للمفحوص، وإنما يستخدمها الأخصائي النفسي لمساعدته على الوصول إلى وصف لفظي أو كيفي للمفحوص، كما في حالة اختبارات الشخصية، ويمكن استخدام كلا المصطلحين على نحو تبادلي كل منهما مكان الآخر في الحالات

التي لا يعبر فيها عن مقدار أو كم الصفة المقاسة، بأي وحدات فيزيائية، فيمكن استخدام أحدهما مكان الآخر مثل مقياس التحصيل الدراسي أو اختبار التحصيل الدراسي أو مقياس الذكاء أو اختبار الذكاء.

وعلى الرغم من هذا التداخل فإنه يمكن القول أن معظم الاختبارات أدوات قياس، ومعظم المقاييس النفسية يمكن أن تستخدم كاختبارات، وذلك عندما يتم التعبير عن درجة وكم الصفة المقاسة بالأرقام.

والاختبارات في جوهرها أدوات للدراسة العلمية للفروق الفردية، وهي بذلك تسعى إلى المقارنة بين الأفراد في ضوء معيار يتحدد من خلال الخصائص الواقعية للأداء الفعلي للأفراد، ويستخدم المتوسط الحسابي لدرجات عينة التقنين معياراً لوصف الأداء العادي في الاختبار، وبالتالي تتحدد الأوضاع النسبية للأفراد، وقد تكون المقارنة داخل الأفراد في ضوء مستوى أو محك.

والمستوى يختلف عن المعيار، في أن الأداء يتحدد في ضوء ما يجب أن يكون عليه الأداء، وليس ما هو عليه بالفعل، مثلما نجده في نظم الامتحانات المعتادة، حين نقارن درجات التلاميذ بنظام النهايات الصغرى والكبرى، أو حين تحدد تقديرات النجاح فيها قبلياً كما في الجامعة " في صورة ضعيف ومقبول وجيد وجيد جداً وممتاز " في ضوء نسب مئوية من النهاية العظمى للمادة الدراسية.

أما المحك فهو أساس خارجي مستقل للحكم على الأداء في الاختبار، وينصب على ما أتقنه الفرد أو أتقنته الجماعة ومستوى أدائهم.

وتعتبر الاختبارات أو المقاييس، وسيلة من الوسائل المهمة التي يعّول عليها الباحث في مجال علوم التربية، وعلم النفس، وعلم الاجتماع، والاختبارات عبارة عن جمع معلومات ما، على عدد من الأشخاص لإعطاء كل منهم درجة تدل على مدى تحصيله أو مدى اتصافه بصفة نفسية خاصة، أو درجة اقتناعه برأي اجتماعي معين أو درجة تعصبه لجهة من الجهات في قياس وتقويم قدراته، وهذه الكفاءة لا تأتى إلا من خلال إعداد اختبارات نموذجية وفاعلة تخلو من الملاحظات التي كثيرا ما نجدها في أسئلة الاختبارات التي يقوم بعض الاشخاص المبتدئين بإعدادها، ومن خلال هذه الخطوات التي قد تساعدنا في بناء اختبارات جيدة تحقق الغرض المطلوب.

فالعلم نظام مفتوح نتائجه متداولة مما يستدعي أن تكون الإجراءات والاستنتاجات العلمية خالية من الغموض، بحيث يستطيع علماء مختلفون تكرارها، والتحقق من صدقها، وهذا لا يتأتى إلا إذا استندت هذه الإجراءات والاستنتاجات إلى أساليب قياس موضوعية.

فالاختبار أداة مهمة يتيح للباحث الحصول على بروفايل من الدرجات، أو المعلومات عن الفرد:مثل أكتشاف قدراته وإمكانياته، وتحديد نقاط القوة أو الضعف لديه. غير أنه لا ينبغي الوقوف على المعنى المجرد للأعداد الإحصائية (الكم) في نتائج الاختبار، وهي أن سلوك المفحوص على نحو معين هو الذي حدد حصوله على هذه الدرجات المعينة، أي ينبغي الوقوف على كيفية أداء المفحوص، وأبعاده الذاتية والموضوعية، وظروفه الخاصة في الموقف المعين.

وتستهدف الاختبارات، بالإضافة إلى اكتشاف قدرات الفرد واتجاهاته، التشخيص والتنبؤ. فمن خلال تشخيص الحالة يمكن التوقع لإتجاه سلوك الفرد في المواقف المختلفة. كما يوضح التقييم في تقرير الوسائل الملائمة، لتخفيف العوامل التي أدت إلى حدوث " الحالة " ورسم خطة علاجية مناسبة. ومن ثم كانت لها أهمية كبرى في عملية رسم استراتيجية التوجيه أو العلاج أي التقويم Evaluation وهذا يعني الدقة في التشخيص والتقييم من أجل تحقيق التوافق الذاتي للفرد. ومن هنا فالاختبارات أو المقاييس تساعد السيكولوجي على إتخاذ القرار.

وواقع الأمر أنه قد يصطدم الأخصائي النفسي أو الباحث بوجه عام في كثير من الأحيان بعدم وجود مقياس مناسب للصفة أو السمة أو الخاصية المراد قياسها، أو حتى لا يناسب أفراد عينته، ومن ثم يصبح لزاما عليه أن يقوم بتصميم مقياس يقيس تلك السمة أو الصفة أو الخاصية التي تتناسب مع أفراد عينته.

وتعتمد عملية تصميم المقاييس في المقام الأول على القيام بعدة خطوات متسلسلة تؤدي في النهاية إلى تجنب كثير من الأخطاء وتتيح إمكانية إعداد مقياس جيد يعتمد عليه في المجال المراد قياسه ، وهي تحتاج تدريبا خاصا نظرا لما تستوجبه من توفر أساس نظري وعملي يعين على القيام بها على الوجه الأمثل.(صديق، 2005: 82)

والقياس مبني على أساس نظرية الفروق الفردية، حيث أن جميع السمات والقدرات موجودة لدى الأفراد ولكن ليس بنفس الدرجة وإنما نسبية وان أي

سمة أو قدرة تتوزع توزيعا طبيعيا، وفي مجال التربية ظهر القياس لحاجة المربين لمعرفة التقدم والتأخر في تحصيل الطلبة والحاجة إلى معرفة نواحي القوة والضعف لديهم، وقد بدأ ذلك بالملاحظة الذاتية والآراء الشخصية، ثم مر بمرحلة ظهور الاختبارات الشفوية فالإنشائية فالموضوعية ومن ثم الأدائية، وأصبحت عملية القياس اليوم جزءا رئيسيا في أي نظام تعليمي، وذلك لأهمية قياس قدرات الطلبة السلوكية والتعليمية، حيث أصبح اليوم من السهولة بمكان اكتشاف الأفراد من ذوي القدرات العقلية المختلفة، وتصنيفهم حسب مستويات الذكاء، لتقديم خطط علاجية لهم، لرفع مستوى قدراتهم في المجالات التحصيلية والتعليمية المتعددة.

ويهدف استخدام الاختبار إلى الحصول على بيانات أو معلومات مهمة عن شخصية المفحوص، وقدراتها، وإمكانياتها، واستعداداتها، واتجاهاتها ودينامياتها. وتعدّ هذه المعلومات من المعطيات الأساسية التي يبني على أساسها الاخصائي استنتاجاته وتشخيصه للحالة، مما يجعل تطبيق الاختبارات ضرورة أساسية لأية ممارسة تقويمية نفسية أو تربوية.

فالقياس هو التقدير الكمي للأشياء وفق معايير محددة، وذلك انطلاقا من القاعدة التي تقول: أن كل شيء يوجد بمقدار، وأن كل مقدار يمكن قياسه، وتتضمن فكرة القياس عادة فكرة المقارنة أيضا، فنحن عندما نقيس، نقوم ضمنيا بعملية المقارنة. والقياس نوعان:

1- ـ قياس مباشر: حيث نقيس الصفة مباشرة كالطول والحجم والوزن، وهذا يتعلق بالنواحي الفيزيائية.

2- قياس غير مباشر: كالظواهر النفسية والتربوية المختلفة، وهنا لا نستطيع قياس الصفة مباشرة ولكن نقيسها بواسطة الآثار المترتبة عليها كالذكاء، والانتباه، والإيمان، والعواطف، والشعور، والاتجاهات.... إلخ.

نشأة الاختبارات والمقاييس:

مرت عملية القياس في علم النفس والتربية بادوار ومجالات متعددة حيث بدأت بالتفسير البدائي، ثم انتقلت إلى المقاييس الحديثة، والتي نراها اليوم كاختبارات الذكاء والقدرات واختبارات التحصيل وغيرها، وكان يعتمد على القياس في مرحلة معينة على النواحي الحسية والحركية، ثم انتقلت إلى العمليات العقلية العليا كالتجريد والتفكير، وقد حدث تطور آخر في حركة القياس، حيث تأثر العلماء بالدراسات العلمية في العلوم الطبيعية والفسيولوجية، حيث كان العلماء يؤمنون في أواخر القرن التاسع عشر أن الناس يختلفون فيما بينهم في قدرتهم على تمييز المثيرات الحسية المتقاربة كالمثيرات اللمسية والصوتية والضوئية، بسبب قدرة الفرد على تركيز الانتباه، ومدى نضج جهازه العصبي، إضافة إلى ارتباط الذكاء بالقدرة السمعية والقدرة على تمييز الأشياء المتقاربة، حتى ظهرت نظريات الذكاء المتعددة والمتطورة من حين لآخر وبناءً عليها ظهرت الاختبارات المختلفة، والتي تقيس سمة أو خاصية ما لكل فرد، وفي أي مرحلة عمرية يمر بها.

وشعر المربون منذ عهود بعيدة بضرورة قياس تحصيل تلاميذهم، ومعرفة نواحي الضعف والقوة لديهم، وللتأكد من صلاحية طرق التدريس والوسائل المستخدمة لتعليمهم، وقد احتلت الملاحظة الذاتية والآراء الشخصية دوراً كبيراً عبر التاريخ، إلا أن الاختبارات مرت بخمس مراحل على النحو التالي:

1) المرحلة الأولى: مرحلة الاختبارات التحريرية: تشير كتابات الصينيين القدماء إلى أن الاختبارات التحصيلية التحريرية كانت معروفة لديهم، وكانت على درجة عالية من الصعوبة، حيث كانت تستغرق الإجابة عنها 24 ساعة أحياناً. ويبدو أن هذا النوع كان معروفاً في المجتمع اليوناني القديم، وفي المجتمع الروماني، واستمر هذا النوع في المدارس حتى بداية العصور الوسطى.
2) المرحلة الثانية: مرحلة الاختبارات الشفهية: عندما عم أوربا الظلام نالت الاختبارات التحصيليه نصيبها منه، فغابت عن مسرح التربية وحل محلها الاختبارات الشفهية, واستمرت الأداة الوحيدة للتقويم حتى العام 1850، وقد انتشرت في أمريكا مع أنها أكثر بلاد العالم اهتماماً بالاختبارات والمقاييس.
3) المرحلة الثالثة: مرحلة الاختبارات التحريرية المقالية: يعتبر عام 1850م تاريخ ميلاد الاختبارات التحريرية وقد استخدمت لاختبار الطلاب بالجامعات الأمريكية، ثم ما لبثت أن أصبحت أداة لقياس التحصيل في المدارس بمختلف مراحلها.

4) المرحلة الرابعة: مرحلة الاختبارات الموضوعية: كانت بمثابة رد فعل على الاتهامات الموجهة للاختبارات المقالية، فجاءت بصفات عديدة تبعدها عن الذاتية،وتجعل الثقة بها أكبر، وهي تستخدم بشكل واسع في المدارس.

5) المرحلة الخامسة: مرحلة الاختبارات الموضوعية المقننة: تعد تطوير للمرحلة السابقة حيث ظهرت الحاجة إلى ضبط وتقنين الإجراءات والتعليمات التي تعطى للمفحوصين، فنشأت فكرة التقنين على المستويات المختلفة، وهذا ما دعا إلى فكرة تقنين المعايير، ويوجد اليوم في الكثير من بلاد العالم الصناعي العديد من الاختبارات الموضوعية المقننة في مختلف المواد الدراسية.

وعلى العموم فإن الاختبارات التي هي أدوات عملية التقويم قد وصلت اليوم إلى مرحلة هامة جعلت من عملية التقويم أحد الأركان الهامة للعملية التربوية.

طبيعة القياس التربوي والنفسي:

لكي تتضح طبيعة القياس التربوي والنفسي، ربما يكون من المفيد إلقاء مزيد من الضوء على القياس في العلوم الطبيعية، فالقياس الفيزيائي، يمكن إجراؤه بطرق مباشرة، حيث إن هناك عمليات وقواعد مقننة متفق عليها لذلك، كما أن هناك موازين Scales معرفة تعريفاً جيداً يتم على أساسها تفسير نتائج القياس.

فعند قياس طول لوح خشب مثلا، ما علينا إلا نأتي بالمسطرة المدرجة، ونجعل إحدى حافتي لوح خشب مقابلة للنقطة المناظرة لصفر التدريج، ونقرأ الرقم المقابل للحافة الأخرى، وبهذه العمليات البسيطة نستطيع معرفة طول لوح

خشب بإحدى الوحدات الخطية المعيارية لقياس الطول مثل السنتيمتر أو البوصة.

غير أن هناك ظواهر فيزيائية أكثر تعقيدا، مثل السمع والبصر وما شابههما، تتطلب إجراءات تفصيلية للقواعد أو الطرق التي تتبع في قياسها. كما أن هناك ظواهر أخرى يصعب قياسها قياساً مباشراً، مثل الصوت والضوء والحرارة والكهرباء وغيرها نظرا لأنها غير مرئية. ولكن هل لا بد من رؤية الظاهرة لكي نتمكن من قياسها؟ بالطبع لا، فعلماء الفيزياء استطاعوا قياس هذه الظواهر الطبيعية وغيرها، على الرغم من أنها غير مرئية أو محسوسة، بل إنهم تمكنوا من قياسها بدقة متناهية بواسطة أدوات قياس متنوعة، فنحن لا نقيس الظاهرة في حد ذاتها، وإنما نقيس أثرها في الأجسام أو الأشياء بواسطة أدوات قياس أعدت لهذا الغرض. فالصوت مثلا يحدث اهتزازات وذبذبات في الأثير على شكل موجات، وتختلف شدة وتكرار الموجات باختلاف نوع الصوت الصادر من مصدر معين، وأداة القياس أو الجهاز المناسب يقيس هذه التأثيرات، ولا يقيس الصوت في حد ذاته، ولكننا نستخدم المقدار الكمي الناتج عن قياس الأثر في الاستدلال على ظاهرة الصوت.

وقد حذا علماء النفس والتربية حذو العلوم الطبيعية، بل إنهم أكثر اهتماما بمنطق القياس والأساليب الكمية من علماء العلوم الطبيعية، ويرجع سبب ذلك إلى أن الظواهر السلوكية ظواهر تتميز بالتعقد والتشابك وتعدد المتغيرات، وليست ببساطة الظواهر الفيزيائية، فعالم العلوم يستطيع أن يثبت زهرة ويقوم

بدراستها، وعالم الكيمياء يمكنه أن يتعامل مع المواد الكيميائية كما يشاء، وكذلك عالم الفيزياء.

أما عالم النفس فمن الصعب عليه إخضاع الإنسان لتجاربه المعملية، وتثبيت جميع المتغيرات التي يود عزل أثرها.

كما أن كثيراً من الظواهر السلوكية يصعب قياسهاعلى موازين مدرجة تدريجا معيارياً متفق عليه كما هو الحال في موازين العلوم الطبيعية، إذ من الصعب قياس سمات شخصية الفرد، مثل العدوانية أو الانطوائية، أو الإقدام أو التسلطية، أو العاطفة، بنفس القدر من الدقة كما في حالة قياس درجة حرارة سائل أو شدة تيار كهربائي.

إلا أن هذا لم يقف حائلاً دون استمرار علماء القياس التربوي والنفسي في مواجهة هذه العقبات، وذلك باتباع خطوات مماثلة للخطوات المتبعة في القياسي الفيزيائي، وابتكار طرق وأساليب وأدوات قياس متنوعة تتميز بخصائص محددة سوف نناقشها فيما بعد.

ولتوضيح ذلك نفترض أننا أردنا قياس "معدل القراءة" لدى تلاميذ أحد صفوف مرحلة التعليم الأساسي، فهنا يجب البدء بتعريف المقصود من الخاصية المراد قياسها، وهي في هذه الحالة "معدل القراءة".

إذ ربما نعرفها بأنها السرعة في القراءة، أي عدد الكلمات التي يقرأها التلميذ في الدقيقة. وتصبح المشكلة الآن هي تحديد مجموعة من الإجراءات أو القواعد لقياس سرعة القراءة.

فربما نلجأ إلى أبسط الطرق وهي أن نطلب من كل تلميذ أن يقرأ قطعة نثرية معينة في مدة زمنية محدودة، ثم نحصي عدد الكلمات التي أمكنه قراءتها في هذه المدة، ونقسم هذا العدد على الزمن المستغرق لنحصل بذلك على قياس لعدد الكلمات المقروءة في الدقيقة، وبهذا نكون قد استخدمنا أسلوباً مماثلاً لأسلوب قياس الطول الذي أشرنا إليه.

ولكننا هنا يجب أن نأخذ بعين الاعتبار مجموعة أخرى من العوامل من بينها صعوبة المقروء، إذ إن هذه الصعوبة تؤثر بلا شك في معدل القراءة، ولذلك فإن عدم معرفتنا درجة صعوبة المقروء يجعل تفسير نتيجة قياس معدل القراءة غير ذى معنى.

والعامل الثاني هو المدة الزمنية المطلوبة للقراءة، فإذا كانت هذه المدة طويلة سوف يشعر التلميذ بالتعب أو الملل ويقل بذلك معدل قراءته. والعامل الثالث هو شروط الموقف الاختباري، فمثلاً إذا كانت الحجرة غير جيدة الإضاءة أو التهوية أو تحيط بها الضوضاء، فإننا نتوقع أن ينخفض معدل القراءة.

كذلك يجب أن نحدد تحديدا دقيقا المقصود بكلمة "القراءة ". وخلاصة ذلك أنه على الرغم من أوجه التشابه بين عمليات وإجراءات القياس الفيزيائي والقياس التربوي والنفسي، إلا أنه يوجد عدد أكبر من المتغيرات التي تؤثر في القياس

التربوي والنفسي يجب ضبطها أو التحكم فيها، إذا أردنا أن نجرى عملية القياس بصورة دقيقة.

القياس النفسي ودوره كأداة تشخيصية:

1- القياس النفسي أداة للحصول على بيانات من سلوك الفرد في موقف معين، وبذلك يمكن جمع بيانات عن هذا السلوك في أسلوب منظم.

فالمادة التي تجمع عن طريق الاختبار هي نتاج مباشر لسلوك الفرد المفحوص. غير أن البيانات التي نحصل عليها من الاختبارات لن يكون لها قيمة إذا اقتصرت دلالتها على السلوك في الموقف المعين دون إمكانية التعميم منها إلى مواقف الحياة. وتلك هي القيمة الكبرى المنسوبة للاختبارات وهي إمكانية التنبؤ بسلوك الفرد.

2- كما أن الاختبارات تمدنا بمعطيات، قد لا يكون المفحوص واعياً بها، أو قد يكون غير قادراً عن التعبير عنها، ومن ذلك التحريفات المبهمة في التفكير والتصور.

وتفيد البيانات الكمية والنوعية في أنها تحدد درجة الفرد، أو مستواه العقلي بالنسبة للمجموعة التي ينتمي إليها... كما انها تلقي الضوء على معنى الدرجة التي حصل المفحوص عليها في الاختبار، ومن ثم، تزداد إمكانية تحديد ظروف التنبؤ الذي يستخرج من تطبيق الاختبار، فضلاً عن ان هذه البيانات تسمح

للفاحص في تحديد الإطار المرجعي للمفحوص ومعرفة قدراته، وحاجاته وأهدافه.

إذاً الاختبار النفسي هو أداة أساسية من أدوات التشخيص النفسي، يستهدف الكشف عن قدرات وإمكانيات الفرد العقلية وكذلك ميوله واتجاهاته المختلفة.

3- الاختبار كأداة للتشخيص النفسي:

تعتبر الاختبارات بأنواعها المختلفة والمتنوعة من أهم الأدوات المستخدمة في علم النفس للتوصل إلى فهم أعمق، ومعرفة أكثر شمولاً لشخصية الفرد، من حيث قدراته العقلية ومشكلاته الإنفعالية.

ويمكن تحديد دور الاختبارات في الأهداف التالية:

- تقييم قدرات الفرد وإمكانياته، من حيث ذكائه العام، وقدراته العقلية الخاصة (اختبارات الذكاء)
- كشف الجوانب المختلفة من شخصية الفرد، مشاعره، أفكاره، ورغباته، واتجاهاته (اختبارات الشخصية).
- تقييم ديناميات السلوك لدى الفرد، وكشف الدوافع الواعية أو اللاوعية التي تحرك هذا السلوك، أي كشف أنواع الصراعات التي يمكن الكشف عنها بالاختبارات الإسقاطية.
- تشخيص المرض أو الإضطراب النفسي أو العقلي الذي يعاني منه الفرد.

" فالتشخيص هو التقييم العلمي الشامل لحالة الفرد، من حيث قدراته وإمكانياته والمشكلات التي يواجهها ويعاني منها ".

وعلى هذا الأساس يعتبر التشخيص من الطرق والأهداف لتخطيط برنامج علاجي شامل، كما يتضمن التشخيص التنبؤ والإفتراضات المتوقعة للحالة المعينة.

تصنيف ادوات القياس:

نظراً لتعقد وتشابك الظواهر التربوية والنفسية، وتعدد التطبيقات الميدانية للقياس في مجالات انتقاء الأفراد لمختلف الوظائف والأعمال والمهام وتسكينهم وتصنيفهم، وفي التوجيه والإرشاد النفسي، وفي تقويم الأفراد والبرامج، والمؤسسات وغير ذلك، فقد تعددت وتنوعت أدوات القياس والتقويم بحيث تناسب طبيعة الظاهرة المراد قياسها وما تشتمل عليه من متغيرات، لذلك أصبح في متناول الباحثين والممارسين العديد من أدوات القياس التي تمكنهم أن ينتقوا من بينها أكثرها ملاءمة للأفراد الذين ستطبق عليهم، ولطبيعة السمة المعينة. فأدوات قياس الذكاء والاستعدادات والتحصيل مثلا تختلف عن أدوات قياس القلق والاكتئاب، أو قياس الميول والاتجاهات،أو الجوانب النفسية الحركية. ومما هو جدير بالذكر أنه لا يوجد تصنيف وحيد لهذه الأدوات، وإنما يمكن تصنيفها وفقاً للعديد من الأبعاد.

تصنيفات ادوات القياس الممكنة:

1- أدوات قياس الجوانب المعرفية في مقابل أدوات قياس الجوانب الوجدانية:

فالجوانب المعرفية للشخصية المتمثلة في القدرات العقلية والتحصيل، تتطلب أدوات قياس يحاول فيها الفرد المختبر بذل أقصى جهده، ليحصل على درجة تعبر تعبيراً صادقاً عن قدرته الفعلية. فالاختبار الذى يقيس الاستعداد اللغوي مثلاً يتطلب من الفرد أن يجيب عن أكبر عدد من المفردات، التي تكون عادة من نوع الاختيار من متعدد في زمن محدد، أما الجوانب الوجدانية والسمات المزاجية للشخصية فإنها تتطلب أدوات لقياس سلوك الفرد الذي يميزه عن غيره من الأفراد، أى سلوكه المعتاد في المواقف الانفعالية مثلاً. وهذه المقاييس تشتمل على عبارات ليس لها استجابة نمطية محددة، وإنما يستجيب الفرد بما يتفق مع خصائصه وسلوكه اليومي المعتاد في مواقف مشابهة للموقف الذى تشير إليه العبارة. ومن أمثلة ذلك: أحب تنسيق الزهور، أو أحبذ الاختلاط في الجامعة، أو أعتقد أن الناس لا تحبني، أو تنتابنى وساوس كثيرة،وهكذا. ونظراً لأن هذه العبارات أو التساؤلات ليس لها إجابة صحيحة أو إجابة خطأ، وإنما تختلف الإجابة باختلاف وجهة نظر أو مشاعر أو معتقدات الفرد، فإن قياس هذه الجوانب يكون أكثر صعوبة من قياس الجوانب المعرفية، حيث يتطلب أدوات قياس تعتمد على أساليب الملاحظة الموضوعية، والتقارير الذاتية، والاستبيانات وغيرها من أساليب التقييم الشخصى.

2- أدوات قياس جماعية في مقابل أدوات قياس فردية:

فأدوات القياس الجماعية تطبق على مجموعة كبيرة من الأفراد في وقت واحد، بينما أدوات القياس الفردية تصمم بحيث تطبق على فرد واحد في وقت واحد، فاختبارات الذكاء الجماعية يمكن أن تطبق على تلاميذ فصل مدرسى في آن واحد، أو على المتقدمين للالتحاق بالخدمة العسكرية، وكذلك اختبارات التحصيل واختبارات الالتحاق بالكليات وغيرها. غير أن هناك مواقف تتطلب استخدام أدوات قياس فردية مثل: مواقف الإرشاد والتوجيه التربوي والمهني أو قياس ذكاء الأطفال، وتطبق هذه الأدوات بواسطة أخصائي متمرس ومدرب تدريباً جيداً على استخدامها، حيث يقوم بملاحظة توجه الفرد المستجيب وأداؤه، وكذلك يستطيع أن يتحاور معه حول استجاباته لكي يتفهمها الأخصائي جيداً، أى أن أدوات القياس الفردية تسمح بقدر كبير من التفاعل بين الفاحص والمفحوص، وتكوين ألفة بينهما، وبذلك تقدم فرصاً ثرية للملاحظات الإكلينيكية. وهناك اختبارات فردية متعددة لقياس الذكاء والشخصية.

ومما لا شك فيه أن أدوات القياس الفردية يستغرق تطبيقها وقتاً طويلاً، وتتطلب غرفة خاصة مجهزة تجهيزاً مناسباً لتطبيق الاختبار في ظروف ملائمة، لذلك فإن هذه الاختبارات الفردية لا تستخدم إلا إذا كان هناك ضرورة لاتخاذ قرار مهم يتعلق بفرد ما، كأن يترتب على هذا القرار نقل طفل من الفصول المعتادة إلى فصول المتخلفين عقلياً مثلاً، أو تشخيص الاضطرابات النفسية.

3- أدوات قياس بمثابة عينة في مقابل أدوات قياس بمثابة مؤشر:

تعتمد بعض أدوات القياس وبخاصة اختبارات التحصيل في بنائها على الانتقاء العشوائى لعينة من المفردات التي تقيس هدفاً أو أهدافاً تعليمية معينة بحيث تمثل هذه المفردات جميع المفردات التي يمكن أن تقيس هذا الهدف أو هذه الأهداف، ومن ثم يمكن الاستدلال على تحقق الهدف أو الأهداف من هذه العينة من المفردات الاختبارية، فمثلا اذا أراد معلم أن يقيس مهارة الطالب في تعريف معنى الرموز الكيميائية، فإنه ينتقى عشوائياً عينة ممثلة لهذه الرموز ويختبر الطالب بها، ومن هذه العينة الصغيرة يستطيع المعلم الاستدلال على تحقق هذا الهدف لدى الطالب، أى أن الاختبار يقيس عينة سلوكية Sample Test، أما إذا استخدم اختباراً معينا للتمييز بين مجموعتين من الأفراد في سمة معينة تختلف عما يقيسه الاختبار بالفعل فإن الاختبار يكون بمثابة مؤشر Sign للتمييز بين المجموعتين في هذه السمة.

فإذا اشتمل اختبار عقلى معين على مفردات تقيس قدرة التلميذ على تذكر مجموعة من الأرقام في تسلسل معين، وتوصلت الدراسات إلى أن عدم قدرة الفرد على ذلك تميز بين المتخلفين عقلياً والعاديين، ففى هذه الحالة لا يصبح قياس تذكر الأرقام هو الهدف من تطبيق الاختبار، وإنما يستفاد من الاختبار كمؤشر لأغراض معينة.

4- أدوات قياس لفظية في مقابل أدوات قياس الأداء:

فمعظم أدوات القياس تتطلب استخدام اللغة سواء في تعليمات المقياس أو مفرداته، لذلك تسمى أدوات قياس لفظية Verbal Tests مثل معظم اختبارات الجوانب المعرفية والوجدانية. ولكن هناك أنواع من أدوات القياس تستخدم مع الأطفال الصغار والأميين والمعاقين لغوياً، تشتمل على أشكال أو رسوم أو صور أو متاهات أو مواد عيانية، فهذه تسمى أدوات قياس غير لفظية Non-Verbal، إلا أن الفاحص يقرأها جهرياً. ولا تتطلب عادة أدوات قياس الأداء إجابات تحريرية لأسئلة معينة، وإنما تتطلب أداء مطلب يتعلق بالمثيرات العينية المعطاة، ومثال ذلك بعض الاختبارات الفردية لقياس الذكاء والاستعدادات المكانية والميكانيكية، ومقاييس الجوانب النفسية الحركية.

ومن الجدير بالذكر أن كثيراً من الاختبارات الفردية للذكاء تتضمن أجزاء لفظية وأجزاء أدائية، وبذلك يمكن قياس الجانب غير اللفظي في الذكاء، وكذلك المقارنة بين درجات كل من الجزئين لتعرف جوانب القصور اللغوى لدى الفرد.

5- أدوات قياس تتطلب السرعة في مقابل أدوات قياس تتطلب القوة:

تتباين أدوات القياس في درجة تحديدها لزمن الإجابة. فهناك اختبارات موقوتة، أى يتحدد فيها زمن الإجابة تحديداً دقيقاً ولا يسمح للفرد تخطى هذا الزمن، وعادة يشتمل الاختبار على مفردات سهلة نسبيا، ولكن نظراً لمحدودية الوقت

فإن كثيراً من الأفراد لا يستطيعون إجابة جميع مفردات الاختبار في الزمن المحدد باستثناء الأذكياء.

ونظراً لأن الهدف من هذه الاختبارات قياس سرعة الأداء، فإن تباين درجات الأفراد لا ترجع إلى صعوبة مفرداته، وإنما إلى اختلاف معدل سرعتهم في الإجابة. وكثير من اختبارات الذكاء والاستعدادات تعد من نوع اختبارات السرعة Speed Tests. أما أدوات القياس التي تهتم بما يستطيع الفرد أداءه بغض النظر عن سرعته، فإنها تسمى اختبارات القوة Power Tests، وهذه الاختبارات تصمم لقياس قدرة الفرد على حل مشكلات تتميز بالصعوبة بغض النظر عن الزمن المطلوب للإجابة. ولعل قياس تحصيل الطلاب يعتمد على هذا النوع من الاختبارات، حيث يكون التركيز على تقييم ما حققه الطالب من معارف ومهارات، لذلك يعطى فرصة كافية للتفكير والإجابة عن جميع المفردات قدر إمكانه.

والحقيقة أنه لا توجد حدود قاطعة بين نوعى الاختبارات، إذ يمكن أن يكون الاختبار سرعة لمجموعة معينة، وقوة لمجموعة آخرى. فمثلاً اختبار فهم معانى الكلمات يعد اختبار سرعة لطلاب المرحلة الثانوية أما إذا كانت مفرداته سهلة ومناسبة لتلاميذ المرحلة الابتدائية، فإنه يعد بمثابة اختبار قوة لهؤلاء التلاميذ. لذلك ينبغى مراعاة التحديد الدقيق لمثل هذه الاختبارات التي تعتمد جزئياً على السرعة عند إعدادها.

6- أدوات قياس موضوعية في مقابل أدوات قياس ذاتية:

فالموضوعية Objectivity والذاتية Subjectivity تتعلقان بطريقة تصحيح الاختبارات والمقاييس التربوية والنفسية وليس بنوع الاختبار أو المقياس.

فقد سبق أن أوضحنا مفهوم الموضوعية كأحد مكونات تعريف الاختبار، وذكرنا أنه يعنى عدم تأثر تصحيح الاختبار بالأحكام الذاتية للقائمين بها. فالاختبارات التي تشتمل على مفردات اختيار من متعدد مثلا يتسم تصحيحها بالموضوعية نظراً لوجود مفتاح تصحيح يطبق على أوراق الإجابات، بل ويمكن تصحيحها بواسطة الحاسوب، لذلك لا تختلف درجة الفرد باختلاف القائمين بعملية التصحيح.

ولكن يمكن أن تختلف الدرجة التي يحصل عليها الفرد في اختبار يشتمل على أسئلة مقالية أو أى نوع آخر من الأسئلة التي تتطلب إجابة مفتوحة باختلاف المحكمين أو القائمين بعملية التقدير، حيث إن عنصر الذاتية يؤدي إلى هذا الاختلاف. ويمكن تلافي ذلك بالعناية بتحديد المطلب الذى يصوغه السؤال تحديدا دقيقاً لا يشوبه الغموض، وكذلك الاتفاق على محكات أو عناصر أساسية ينبغي توافرها في الإجابة للاسترشاد بها في تقدير الدرجات مما يقلل من الذاتية، ويزيد من اتساق الدرجات، وبذلك يكون التصحيح أكثر موضوعية. وعموما، فإن تقنين عمليات تطبيق الاختبار أو المقياس وتعليماته وطريقة تصحيحه يمكن أن تسهم في تحقيق الموضوعية المرجوة بغض النظر عن نوع أداة القياس المستخدمة.

7- أدوات قياس يعدها المعلم في مقابل أدوات قياس مقننة:

فالمعلم يهتم بقياس مدى تحقيق تلاميذه للأهداف التعليمية للمجالات الدراسية ليس فقط لوضع تقديرات لكل منهم في نهاية مدة دراسية معينة، وإنما لتبين جوانب القوة والضعف في تحصيلهم لكي يتسنى له اقتراح أساليب علاج مناسبة للطلاب الذين يجدون صعوبة في تحقيق هذه الأهداف.

لذلك فإنه يقوم بإعداد اختبارات يكون لها وظيفة تعليمية مهمة في تحسين دافعية الطلاب وحثهم على بذل المزيد من الجهد لإثراء تعلمهم، وعادة تكون هذه الاختبارات من نوع الاختبارات مرجعية المحك، غير أن هذا يعتمد على تمكن المعلم من المهارات الأساسية في بناء مثل هذه الاختبارات الصفية.

ولكن المعلم يهتم في بعض الأحيان بمقارنة تحصيل طلاب فصل أو فرقة معينة في مجال أو أكثر من المجالات الدراسية بمجموعة خارجية ممثلة لمجتمع هؤلاء الطلاب ولتكن مجموعة طلاب الفرقة نفسها على المستوى الوطنى التي تسمى الجماعة المرجعية Norm Group، فإنه عندئذ ينبغى أن يعتمد على اختبارات تسمى الاختبارات المقننة Standardized Tests. وبقوم ببناء هذه الاختبارات خبراء في القياس التربوي والمناهج والمعلمين، وتجرى عملية التقنين عن طريق تطبيق الاختبار وتصحيحه في ظروف موحدة ومنظمة على الجماعة المرجعية، والتوصل إلى معايير Norms يمكن في ضوئها مقارنة نتائج فصول فرقة معينة أو مدرسة معينة بأقرانهم على المستوى الوطنى. فالدرجة التي يحصل عليها أحد طلاب فرقة معينة في إحدى المدارس في مجال دراسي معين يمكن مقارنتها بمعيار

الجماعة المرجعية الذى تم التوصل إليها لمعرفة المستوى التحصيلى لمجتمع طلاب هـذه الفرقة في المجال نفسه، لذلك فإن مثل هذا الاختبار يكون مـن نـوع الاختبـارات مرجعيـة الجماعة أو المعيارNorm- refrenced test.

ومن الجدير بالذكر أن معظم الاختبارات والمقاييس النفسية تعـد أيضاً مـن هـذا النـوع، نظراً لأنها تعتمـد عـلى مقارنـة أداء الفـرد بـأداء أقرانـه في الجوانـب العقليـة والوجدانيـة والمزاجية. أى أن الاهتمام يكون منصباً على ترتيب الأفراد بعضـهم بالنسـبة للـبعض الآخـر من حيث درجـاتهم في الاختبـار، ويكـون للدرجـة تفسـيراً فقـط في ضـوء معيـار Group Norm يعتمد على متوسط أداء الجماعة التي ينتمى إليها. وتختلف الاختبارات مرجعيـة الجماعة عن الاختبـارات محكيـة المرجـع Criterion Referenced Tests في أن الأخـيرة ترجع إلى محـك إتقـان معـين، كـأن أقـول يـنجح طالـب الجامعـة بتقـدير ممتـاز إذا بلـغ مجموعه التراكمي 90%.

إستخدام أدوات القياس:

تستخدم أدوات القياس التربوي والنفسي في أغراض متعددة، يصعب حصرهـا نظـراً لتنـوع الظـروف والمواقـف، التـي تتطلـب تطبيـق هـذه الأدوات للاسـتناد إلى نتائجهـا في اتخـاذ قرارات.

فالباحــث التربــوي والــنفسي يحتــاج في تصــميم دراســته وتنفيــذها إلى العديــد مــن أدوات القيـاس المناسـبة، التـي تمكنـه مـن التحقـق مـن فـروض دراسـته، سـواء كانـت الدراسـة تجريبيـة أو ارتباطيـه أو وصـفية أو إكلينيكيـة، كـما يحتـاج اليهـا الممارسـون

ومتخذو القرارات في مختلف المجالات التربوية والنفسية مثل المعلم، والاختصاصى النفسي والاجتماعى، والقائم بالعلاج النفسي، والمسؤولون عن انتقاء الأفراد لشغل مختلف أنواع الوظائف والمهن أو تصنيفهم وفقاً لمحكات معينة، والمهتمون بتقويم المشروعات والبرامج والمؤسسات وغيرهم.

مجالات الاختبارات النفسية: تستخدم الاختبارات النفسية في مجالين: المجال الأول قياسي (سيكومتري) يعتمد على الكم في التعامل مع درجات الاختبار، والمجال الثاني تحليلي يعتمد في الغالب على الكيف في التعامل مع الاستجابات. ونشير إلى كل مجال كالآتي: المجال الأول: قياس الشخصية باستخدام الاختبارات السيكومترية:

الاختبارات السيكومترية التي نستخدمها في مجال الإرشاد: اختبار القدرات العقلية، والاستعدادات، واختبارات الشخصية التي تقيس جانباً واحداً مثل الشعور بالذنب والمخاوف أو التي تقيس أكثر من جانب مثل سمات الشخصية، وأنواع الانحرافات النفسية، والاختبارات التحصيلية التي تقيس الحفظ والاستيعاب والتطبيق والتحليل والتركيب والتقويم. وتتسم بالموضوعية والصدق والثبات.

أهداف الاختبارات السيكومترية: يمكن تصنيف أهداف الاختبارات السيكومترية إلى نوعين هما: أ- قياس أقصى الأداء: يتم ذلك من خلال قياس أفضلية الأداء أو سرعة الأداء. ومن أمثلة الاختبارات في هذا المجال: اختبارات الذكاء، واختبارات الاستعداد. ب- قياس الأداء النمطي: يتم ذلك من خلال قياس الخصائص النفسية والعادات والاتجاهات والميول، والتوافق، وهذا النوع

يركـز عـلى مـا إذا كانـت الخاصـية موجـودة أم لا، ويحـدد درجتهـا ودلالتهـا. والاختبارات السيكومترية قد تعتمد على اللغة (الاختبارات اللفظية) أو تعتمد على الرموز والأشكال (الاختبارات غـير اللفظيـة) الـذى يصـلح اسـتخدامها مـع الأطفـال الذين لم تنمو لديهم لغة التعبير بعد. المجـال الثـانى: تحليـل الشخصـية بأسـتخدام الاختبارات الاسقاطية progective test، الاسقاط فى ضوء التحليـل الـنفسى يعـرف بأن ميكانيزم دفاعى يلجأ اليه الشخص ليسقط مشاعره ودوافعه على أشـخاص أو موضـوعات خارجيـة، ويشـمل الاسـقاط العمليـات الشـعورية واللاشـعورية عـلى السواء، بالتـالي يكشـف عـن ميـول واتجاهـات الفـرد الواضـحة. وفي ضـوء نظريـة "الجشطلت" أصبح فى الامكان تكوين تصور كلى عن الشخص من خلال اسـتجابات للمثيرات. وتقوم فكرة الاختبارات الاسقاطية على تقديم مثير غـامض للشـخص ثـم يطلب منه أن يعطيه المضـمون فى ضـوء ادراك لهـذا المثـير وهـذا يعكـس جوانـب شخصيته الديناميكية، أي انه يسقطها على المثير من أفكار ومشاعر ومخاوف التى يعاني منها.

الاستخدامات الرئيسة لأدوات القياس المختلفة:

- انتقاء الأفراد: تستخدم أدوات القياس في غالبية الأحيان في تزويـد المعنيـين بشؤون الأفراد ببيانات مفيدة يمكن الاستناد إليها بقدر محدد من الثقة في اتخاذ قرارات تتعلق بانتقـاء الأفـراد Selection لشـغل وظـائف أو مهـن معينة، أو انتقاء الطلاب للالتحـاق بالكليـات والمعاهـد ومراكـز التـدريب المهني وغيرها، فإلى جانب ما تحدده هذه المؤسسات التربوية من حد أدنى

من المتطلبات، فإن الأماكن الشاغرة فيها تقل عن عدد المتقدمين. فهنا يمكن أن تطبق اختبارات تقيس مدى توفر هذه المتطلبات لدى المتقدمي، وتسمى اختبارات مرجعية المحك Criterion-Referenced Tests، كما يمكن تطبيق اختبارات للتنبؤ بنجاح الطلاب أو رسوبهم خلال دراستهم أو العمل بأقل قدر ممكن من الخطأ. والمقصود بالخطأ هنا قبول طلاب يتبين رسوبهم خلال دراستهم، أو رفض طلاب كان من الممكن نجاحهم لو أتيحت لهم فرصة القبول، وهذه الاختبارات التنبؤية ينبغى ان تقلل احتمال الوقوع في أى من هذين النوعين من الخطأ قدر الإمكان،وتسمى هذه الاختبارات مرجعية الجماعة أو المعيار.Norm-Referenced Tests وهذا يتطلب مراعاة أن يكون هناك تناظر كبير بين المفردات التي تشتمل عليها الاختبارات، التي تستخدم في الانتقاء وبين المتطلبات والمهارات والقدرات المطلوبة للنجاح في الدراسة أو العمل، وإجراء العديد من الدراسات التنبوئية قبل استخدام الاختبارات في هذا الشأن.

- تسكين الأفراد: إن انتقاء الأفراد للالتحاق بالدراسة بإحدى الكليات أو لشغلهم وظيفة أو مهنة معينة يعنى المزاوجة بين الفرد والعمل، أى وضع الشخص المناسب في المكان المناسب. ولكن التسكين Placement يعنى وضع الطالب الذى تم قبوله بإحدى الكليات مثلا في قسم معين يناسبه بحيث يحقق فيه تفوقاً وتميزاً، أو وضع التلميذ الذي تم قبوله في المدرسة في برنامج تعليمى معين يناسبه مثل البرامج العلاجية، أو العادية، أو الإثرائية. وكذلك وضع الفرد الذى تم انتقاؤه في أحد الأقسام المتعلقة بالمهنة أو

الوظيفة التي شغلها. ويختلف التسكين عـن الانتقـاء في أن قـرارات الانتفـاء تتضـمن قبول الفرد أو رفضه، فإذا ما رفض عليه أن يبحث عن مكان آخر، أمـا إذا تـم قبولـه فإن المؤسسة تعمل على تسكينه وفقا لإمكاناته واحتياجاته. فالتلميذ الـذي يقبـل في مدرسة مهنية معينة ينبغى أن تعمل المدرسة على تسكينه في قسـم مناسـب لمهنتـه، أو تقدم لـه برنامجـا تـدريبيا علاجيـا، أو تقـوم بترفيعـه في مسـتوى متقـدم بحسـب المعلومـات التـي تقـدمها اختبـارات التسـكين التـي تسـمى عـادة اختبـارات تحديـد المستوى.

لذلك فإن الاختبار الذى يسـتخدم في اتخـاذ قـرارات الانتقـاء، ربمـا لا يصـلح في اتخـاذ قرارات التسكين، إذ ربما يصلح اختبار يقيس الاستعداد الدراسي في التنبـؤ بالنجـاح في الكليات، ولكنه ربما لا يصلح في تسكين الطالب المقبـول في البرنـامج التعليمـي الاكـثر ملاءمة له.

- تصنيف الأفراد: على الرغم من أوجه الشبه بين التصنيف والتسكين والانتقـاء، إلا أن نيتكـــو (Nitko, 1983) يـــرى أن التصـــنيف Classifiation أكـــثر عموميـــة منهما. فقرارات التصـنيف تتعلـق بتعيـين الأفـراد في أقسـام أو مجموعـات مختلفـة، فعنـدما يـتم انتقـاء الأفـراد للالتحـاق بالخدمـة العسـكرية مـثلاً يقـوم المسـؤولون بتصنيفهم في وحدات عسكرية مختلفة، وتسكينهم في أعمال مناسبة لهم داخل هذه الوحـدات، كـما أن مـرضى الاضـطرابات العقليـة يجـري عـادة تصـنيفهم، وكـذلك الأطفـال مـن ذوى الإعاقـات مثـل الأصـم، والأصـم - الأعمـى، و ضـعاف السـمع،

ومتعدد الإعاقات، وشديد الاضطراب الانفعالى، والمعاق بصرياً وهكـذا. لـذلك يقـترح نيتكو Nitko أن التصنيف يتعلق بتعيين الأفراد في أقسـام أو أعـمال أو بـرامج. فـإذا كانت هذه الأقسام تمثل مستويات مرتبة لنوع معـين مـن الـبرامج أو الأعـمال (مثـل المستويات العلاجية والعادية والإثرائية) فإن هذا التعيين يسمى " تسكين "، وإذا كان أحد هذه الأقسام يمثل الرفض والأخر يمثل القبول، فإن التعيين يسمى " انتقـاء"، أمـا "التصنيف " فيمثل أقساما غير مرتبة بطبيعتها، والقرارات المتعلقة بهما جميعـاً تعـد من نوع القرارات الإدارية.

- التشخيص والعلاج: تستخدم أدوات القياس النفسي والتربوي في تشخيص الاضطرابات النفسـية Diagnosis، وكـذلك في تشـخيص جوانـب القـوة والضـعف لـدى الطـلاب وأصحاب المهن، من أجل اقتراح أساليب علاج مناسبة، فمعلم التعلـيم الأسـاسى مـثلاً يعمل على تشخيص جوانـب الضـعف لـدى تلاميـذه في المهـارات الأساسـية المتعلقـة بـالقراءة والكتابـة والتخاطـب والحسـاب، والمرشـد الـنفسي يقـوم بتشـخيص المجالات السلوكية التي يعاني منها الفرد. فالمعلم يحاول معاونة تلاميـذه في تحسـين أدائهم، والمرشد النفسي يعمل على تقديم برنامج علاجي مناسب للفـرد الـذى يعـاني من مشكلات نفسـية معينـة وهكـذا. ومـن الجـدير بالـذكر أن الاختبـارات التربويـة والنفسية التى تستخدم في اتخاذ قـرارات التشـخيص تختلـف عـن الاختبـارات التـي تسـتخدم في الانتقـاء والتسـكين والتصـنيف في أنهـا تكـون أكـثر تفصـيلاً وعمقـاً، نظـراً لتركيزهـا عـلى جوانـب معينـة في السـمات المـراد قياسـها وتعـرف جوانـب

الضعف فيها، ولذلك تسمى الاختبارات التشخيصية Diagnostic Tests، ويتطلب بناء هذه الاختبارات جهداً كبيراً وخبرة فنية عالية في تقنيات بنائها. (علام، 1995)، وذلك لأهميتها وضرورتها في مختلف المجالات التربوية والنفسية وفي كثير من التطبيقات العلمية.

- الارشاد والتوجيه التربوي والمهني: فأدوات القياس المتنوعة يمكن استخدامها استخداماً وظيفيا فاعلا في عمليات الإرشاد Counseling والتوجيه Guidence التربوي والنفسي. وعلى الرغم من أن قرارات الانتقاء والتسكين والتصنيف تعد من القرارات الإدارية المؤسسية، إلا أن قرارات الإرشاد والتوجيه تتعلق بفرد معين شأنها شأن قرارات التشخيص والعلاج. فقد سبق أن ذكرنا أن المؤسسة تتخذ قراراتها في الانتقاء مثلا بما ينطبق على كثير من الأفراد المتقدمين في ظروف متشابهة، بحيث تكون هناك درجة معينة من الثقة فيما يترتب على هذه القرارات، فاحتمال نجاح الأفراد في دراستهم أو في أعمالهم بدرجة كبيرة هو المحك الذى تستند إليه هذه القرارات،أما القرار المتعلق بفرد معين فيؤثر على هذا الفرد بذاته ولا يؤثر فيما عداه، وعليه أن يتحمل المخاطرة في حالة عدم التزامه بهذا القرار.

فالفرد الذى يسعى إلى المرشد التربوي أو النفسي يود أن يتعرف على إمكاناته واستعداداته، ويحدد مجال العمل أو الدراسة الأكثر ملاءمة له. لذلك فهو يلجأ إليه بمحض إرادته ويتعاون معه تعاوناً مثمراً وبخاصة في المواقف الاختبارية. فالتوجيه والإرشاد التربوي يهتم بتعريف الطالب بما

يمكنه دراسته من مقررات والتخصص الدراسي الذى يناسبه. والتوجيه والإرشاد المهني يهدف لمعاونة الفرد على اختيار مجال عمل يستطيع أن يؤديه أداء متميزاً بما يتفق واستعداداته وسماته الشخصية، مثل مجال الطب أو الهندسة أو الأعمال الحرة أو الفنية أو الحرفية. والتوجيه والإرشاد الشخصي يركز على المشكلات الشخصية، لدى الفرد التي ربما تعكس اضطرابات في الشخصية مثلا أو تكون ناتجة عن سوء تكيف وقتي، مثل سوء تكيف المراهق مع أقرانه الأكثر تفوقاً، أو سوء تكيفه مع الجنس الآخر.

وفي جميع الحالات ربما لا يقتصر المرشد التربوي أو النفسي على استخدام الاختبارات، وإنما يمكنه أن يجمع معلومات ثرية من مصادر متنوعة. إذ ربما يقوم بمقابلات شخصية تفصيلية مع العميل لتعرف وجهات نظره وخلفيته واتجاهاته ورغباته وأهدافه وخصائصه، وربما يلجأ إلى جمع معلومات عن الفرد من مصادر أخرى ذات علاقة به مثل السجل المدرسي، أو آراء أقرانه ومعلميه ووالديه وأصدقائه، وغير ذلك.

وانتقاء الأدوات المناسبة يعتمد على خبرة المرشد النفسي بهذه الأدوات وتمرسه على تطبيقها والإفادة من نتائجها لكي يتخذ قراراً صادقاً فيما يتعلق بالعميل.

وعلى الرغم من أن الانتقاء والإرشاد يستندان إلى عمليات أساسية متشابهة حيث يهتم كلاهما بالتنبؤ بالنجاح، إلا أن الانتقاء يركز بدرجة أكبر على عملية التنبؤ، كما يستندان إلى نتائج أدوات قياس متنوعة على الرغم من

اختلاف الإطار المرجعي لكل منهما، فقرارات الانتقاء تنتهي بقبول فرد له خصائص معينة في دراسة أو مهنة معينة مناسبة لخصائصه من بين أعداد كبيرة من المتقدمين للدراسة أو المهنة نفسها بحيث يتوقع أن يكون هذا الفرد أفضل المتقدمين. وعلى العكس فإن قرارات التوجيه والإرشاد تهدف لتحديد مجال العمل المناسب للفرد بحيث يستثمر فيه إمكاناته وقدراته استثماراً فاعلا. وتتأثر قرارات التوجيه والإرشاد بمدى دقة أدوات القياس في التنبؤ بالأداء في عمل معين، ولذلك يفضل التنبؤ بالأداء في مجال عمل محدد بدلاً من التخصيص، حيث إن الفرد ربما يمكنه أداء أكثر من عمل واحد بفاعلية، فالنجاح في أداء الأعمال يعتمد إلى جانب الاستعدادات على المناخ السائد في العمل، واتجاهات المشرف على العمل، وسياسة المؤسسة، والعلاقات بين الزملاء، وتوفر الأجهزة المناسبة.

الأخطاء الناجمة عن إساءة استخدام الاختبارات والمقاييس النفسية والتربوية:

على الرغم من أن الاختبارات والمقاييس النفسية والتربوية تعد أحد مصادر جمع المعلومات،إلا أنها بما تتميز به من خصائص مهمة، مثل الموضوعية والتقنين والمعايير والثبات والصدق، تعد من أهم هذه المصادر (Cronbach, 1970)، غير أن سوء استخدام هذه الأدوات من قبل أفراد غير متخصصين أو غير متدربين تدريباً عالياً، وغير مطلعين على التطورات المتنامية في مجال القياس والتقويم النفسي والتربوى، يؤدى إلى أضرار بالغة بمن تطبق عليهم هذه الأدوات، ويفقدها فعاليتها، ويحيد بها عن تحقيق الأغراض الأساسية منها، بل

وتصبح مصدراً من مصادر الشك والنقد من مختلف فئات المجتمع (Aiken, 1980)، فالممارسات الخاطئة لاستخدام الاختبارات والمقاييس النفسية لا تؤدى فقط إلى قرارات غير صائبة، وإنما لها آثاراً ضارة بالمفحوص وانعكاسات اجتماعية سيئة.

ولا نتعدى الحقيقة إذا قلنا أن هذه الممارسات الخاطئة كانت سبباً أساسياً في الهجوم الشديد والنقد المستمر للاختبارات والمقاييس النفسية ذاتها.

وأهم هذه الممارسات الخاطئة وما يترتب عليها من مشكلات ما يلى:

- تحيز المسؤولين عن البرامج الاختبارية: فبعض المسؤولين عن البرامج الاختبارية أو التقويمية لا يراعون تباين الأفراد المختبرين في خلفياتهم الثقافية وخصائصهم الشخصية المميزة عند انتقاء الاختبارات والمقاييس النفسية واستخدامها، وهذا يؤثر بلا شك في نتائج الاختبارات، وبالتالى يقلل من شأن هؤلاء الأفراد.
- الثقة التامة في التنبؤ باستخدام الاختبارات: إن الثقة التامة في التنبؤ باستخدام الاختبارات والمقاييس المقننة يعد من الممارسات الخاطئة. فالتنبؤ بالسلوك المستقبلي باستخدام مختلف أنواع الاختبارات والمقاييس ينبغى أن يكون في حدود معينة من الثقة. إذ إن هذا التنبؤ يرتبط بعوامل متعددة تتعلق بالفرد المختبر، والسلوك المراد التنبؤ به، والمدة الزمنية للتنبؤ، وخصائص المقياس المحك الذى يقدر في ضوئه فاعلية التنبؤ، وإذا قلت الثقة

في التنبؤ بدرجة ملحوظة يصبح استخدام هذه الأدوات مضللاً، وبخاصة إذا كان القائم باستخدامها شخص غير مدرب أو قليل الخبرة.

كما أن استخدام الاختبارات في أغراض الانتقاء والتصنيف التي لم يتم التحقق من صدقها أو اتبع في سبيل التحقق من ذلك إجراءات واهنة يعد أيضاً من الممارسات الخاطئة. فمشكلة صدق الاختبارات والمقاييس النفسية تتضمن في واقع الأمر قضايا متعددة، ولا تقتصر فقط على مجرد تقدير خاصية من خصائص الاختبار أو المقياس. فالاستخدام المناسب للاختبارات يتضمن العديد من المسؤوليات التي يعتمد بعضها على درجة ثبات وصدق الاختبار، ويعتمد البعض الآخر على الظروف المختلفة المتعلقة بتطبيقات الاختبار.

❖ الاستخدام الروتيني لدرجات الاختبارات:

فالدرجات المستمدة من الاختبارات تزودنا بقاعدة من البيانات تفيد في تصنيف الأفراد تصنيفاً اعتبارياً. وهذا يعنى أن هذه الدرجات ليست قياسات ثابتة ولكنها قيماً تقديرية لسمات انسانية، وبالتالى فهي تحتوى على أخطاء معاينات مختلفة (1982 Anastasi,) غير أن هذه الاختبارات تستخدم عادة بطريقة روتينية لا تتسم بالمرونة الكافية التي تأخذ مختلف المتغيرات المؤثرة بعين الاعتبار. وهذه الممارسة الخاطئة تؤدي في بعض الأحيان إلى تصنيف التلاميذ مثلاً في مجموعات بحسب قدراتهم، اعتماداً على اختبارات أو مقاييس ربما يصعب تبرير استخدامها، أو استخدام

درجات فاصلة Cut-off Scores في الاختبارات واعتبارها محكات شبه ثابتة لانتقاء الأفراد في مجالات عمل معينة في المؤسسات الصناعية.

❖ افتراض أن الاختبارات العقلية تقيس خصائص موروثة:

فأحياناً يفسر بعض مستخدمى الاختبارات درجاتها على أنها قياسات لخصائص موروثة، ومثال ذلك سمة الذكاء. وهذا التفسير الخاطئ له آثار غير مرغوبة،إذ إنه يؤدي إلى اعتبار أن القدرات الإنسانية ثابتة أو غير قابلة للنمو، بل وربما يؤدي إلى النظر إلى الإنسان نظرة تصنيفية، كأن نقول أن هذا متخلف عقلياً وهذا عبقرى وما شابه ذلك، ويرى بعض النقاد أن هذه نظرة غير إنسانية.

☒ أخطاء ناجمة عن تقديم الاختبارات إلى غير المتخصصين أو إلى المفحوصين:

هناك ممارسات خاطئة أيضاً تنجم عن تقديم نتائج الاختبارات والمقاييس النفسية إلى أفراد غير متخصصين أو إلى المفحوصين. ومن بين هذه الأخطاء ما يلى:

❖ مسايرة التوقعات لنتائج الاختبارات: فقد بينت دراسة روزنثال وجاكوبسون (Rosenthal & Jacobson,1970) أنه إذا أعطى للمعلمين معلومات افتراضية - أى ليست حقيقة - عن قدرات تلاميذهم فإن التحصيل الفعلى لهؤلاء التلاميذ يتفق وتوقعات المعلمين عنهم. فالتلاميذ الذين توقع المعلم أن يكون تحصيلهم منخفضاً كان

تحصيلهم كذلك بالفعل، والعكس كان صحيحاً. لذلك فإن الذين ينتقدون اختبارات الذكاء يرون أن هذه الاختبارات تؤثر في أداء الأفراد تأثيراً غير مرغوب فيه. إذ يرون أن المعلمين مثلاً ربما ينظرون إلى التلاميذ مرتفعي الذكاء نظرة تختلف عن نظرتهم للتلاميذ الأقل ذكاء في ضوء ما يتوقعونه من كل منهم.

❖ تأثر مفهوم الذات لدى الفرد بنتائج الاختبارات: فالأخصائي النفسي غير المتمرس ربما يقدم نتائج الاختبارات والمقاييس النفسية للمفحوص بطريقة تؤثر سلبياً في مفهومه عن ذاته. إذ ربما يؤدي تعميم المفحوص لنتائج الاختبار إلى استنتاج أشياء عن نفسه ليس لها ما يبررها، فالشاب الصغير الذى يحاول جاهداً أن يكون مفهوماً إيجابياً عن ذاته ربما يتأثر هذا المفهوم لديه ويقل مستوى طموحه إذا علم أن درجاته منخفضة في بعض الاختبارات. فالتقارير المقتضبة لنتائج الاختبارات، وما يبنى عليها من توصيات تقدم للأفراد في بعض مواقف الارشاد النفسي، تؤثر بلا شك تأثيراً سلبياً في نظرتهم إلى أنفسهم، وفي تشكيل اتجاهاتهم نحو ذواتهم.

الفصل الثاني

مقدمة:

تعد خطوة تحديد فكرة المقياس ومبررات تصميمه من أهم الخطوات وأولها نظرا لأنها تتيح للقائم بتصميم المقياس الوصول للمداخل والأفكار الرئيسية التي سوف يستند إليها في تصميمه، وتعتمد عملية تصميم المقاييس على القيام بعدة خطوات متسلسلة تؤدي في النهاية إلى تجنب كثير من الأخطاء وتتيح إمكانية إعداد مقياسا جيدا يعتمد عليه في المجال المعني، وهي تحتاج تدريبا خاصا نظرا لما تستوجبه من توافر أساس نظري وعملي يعين على القيام بها على الوجه الأمثل، وتقوم هذه الخطوة بدور الموجه الذي يعين مصمم المقياس خلال الخطوات التالية.

مراحل تصميم الاختبار:

يخضع تصميم الاختبار الجيد لعدد من الاعتبارات والقواعد المنهجية، كما تتطلب فنية استخدام الاختبار دراية واسعة وعميقة بالأساليب المختلفة، ومدى مناسبة كل منهما للأغراض المختلفة. وأساليب تصميم وانتخاب فقرات صادقة وثابتة ومميزة هي أساليب عملية وعلمية معاً. ويتضمن الجانب العلمي فيها المعرفة الدقيقة بكيفية وضع المفردة في الاختبار ومهارة تصميمه، مع حساب قدرته التمييزية ومستوى صعوبته وثباته، ويأتي الجانب الفني في الربط التكاملي بين هذه الخصائص العملية والخصائص اللفظية للفقرة والاعتبارات العلمية التي سيستخدم فيها الاختبار ولا يعد تصميم وإعداد الاختبار جديد هدفاً في حد ذاته، فالباحث لا يلجأ عادة لهذا العمل إلا إذا دعته إلى ذلك حاجة ملحة،

وعندما يتضح له أن الاختبارات المتوفرة لا تفي بغرضه، ودون وجود الدواعي الضرورية لتصميم اختبار جديد، يظل استخدام الاختبارات المتوفرة في المجال أو تطويرها أفضل كثيراً للاستفادة من الدراسات والنتائج التي سبق التوصل إليها.

ويمكننا أن نلاحظ أن هناك اتجاهين بارزين في تصميم الاختبارات هما:

الاتجاه الأول: علمي نظري يهدف إلى تصميم الاختبار وفق إطار نظري محدود، وللإجابة على فروض جيدة الصياغة يقدمها الباحث، ويصمم وينتخب بعناية فائقة فقرات الاختبار وفقاً لمدى قربها من الفروض الأساسية.

الاتجاه الثاني: عملي أو فني وهو يرمي لمجرد تطوير اختبارات جديدة لتكون أداة في يد الممارس أو الأخصائي النفسي بغض النظر عن اعتبارها أداة علمية للبحث.

والاختلاف بين هذين الاتجاهين له تأثيرات مهمة على منهجية تصميم الاختبارات، فمصححوا الاختبارات ممن يأخذون بالمنحنى العملي، ويطورون أساليب خاصة بهم تختلف بقدر ما عن تلك التي يستخدمها الباحث، الذي يصب القدر الأكبر من اهتمامه على صياغة فروض النظرية. ورغم أن صياغة الفروض جوهرية للغاية في تصميم الاختبار، إلا أن عمليات صياغة هذه الفروض تختلف تماماً في الحالتين بين الباحث والفني، فبالنسبة للعالم تعد مشكلة صياغة الفروض أهم خطوات العمل كلها وأصعبها، وهي ما يتطلب منه الأصالة إذ لا بد أن تكون فروضه منهجية إلى تقديم إضافة حقيقية مع اتساقها في الوقت نفسه مع الحقائق القائمة والملاحظات السابقة والتبريرات المنطقية التي

يقدمها الباحث لفروضه، هي التي تميز عمله عن العمل الفني الذي يهدف إلى إجراءاته المهارية في إصابة الهدف وتحقيق الأغراض المباشرة، ورغم ذلك نستطيع أن نجد المزايا والعيوب في كلا الاتجاهين العملي الفني سريع النتائج، وتصميم الاختبارات فيه لا يتطلب وقتاً طويلاً، وقد يبدأ من معلومات خاطئة ولكنه يؤدي لنتائج عملية مفيدة ينتهي بمعلومات محدودة ذات قيمة نظرية ضئيلة. بينما المنحى العلمي النظري يتطلب وقتاً في إعداد الاختبارات، ووقتاً في تنظيم وتحليل النتائج، وقد تكون نتائجه العملية محدودة، ولكنه يبدأ من فروض صلبة وينتهي بمعارف هامة، ويؤدي في أحيان كثيرة إلى الإسهام في تطوير النظرية.

أهداف الاختبار:

أولاً: يتعين قبل أن يبدأ بإعداد الاختبار أن يحدد الهدف اللازم من ذلك الاختبار بل ويحدده بدقة متناهية لما سيترتب على هذا الهدف من نتائج.

فالأهداف التي يوضع من أجلها الاختبار كثيرة ومتباينة، وقد يكون الغرض منه قياس تحصيل الطالب بعد الانتهاء من دراسة أو الانتهاء من وحدة دراسية معينة، وقياس تحصيله، وقد يكون الهدف من الاختبار تشخيصياً لتحديد جوانب التأخر والضعف الدراسي في موضوعات محددة، أو في مواد دراسية معينة كالاختبارات التي يتم تنفيذها في البرامج العلاجية لتحسين مستوى التحصيل عند بعض الطلبة، وما إلى ذلك.

فإذا لم يحدد الباحث أو المعلم هدفه بوضوح سوف لا يكون للنتائج المتمخضة عن ذلك الاختبار أي معنى يذكر.

ثانياً: ـ تحديد الأهداف السلوكية أو الإجرائية، أو ما يعرف بالنواتج التعليمية:

عند إعداد الاختبار، يجب على الباحث أن يركز على خطوة أساسية، وهامة تكمن في تحديد الأهداف السلوكية، أو ما يعرف بنواتج التعلم التي خطط لها قبل أن يبدأ في عملية التدريس. ويتم تحديد نواتج التعلم من خلال صياغة تلك الأهداف صياغة إجرائية محددة وواضحة بعيداً عن الغموض والتعميم.

بعض الإرشادات التي يجب مراعاتها عند صياغة الأهداف الإجرائية:

1) يجب أن تصف عبارة الهدف سلوك المتعلم وأدائه الذي يستدل منه على تحقيق الهدف، ولا تركز على سلوك المعلم.
2) أن تكون عبارة الهدف واضحة الصياغة لا تقبل إلا تفسيراً واحداً.
3) أن يصف أداء قابلاً للملاحظة والقياس.
4) أن يشتمل على الحد الأدنى للأداء المقبول " المعيار الكمي والكيفي ".
5) يجب أن يكون الهدف بسيطاً " غير مركب " بمعنى أن كل عبارة منه تتعلق بعملية واحدة فقط.

أسلوب إعداد الاختبار:

☒ تحديد المجال الذي يراد قياسه ليكون إطار يشتق منه عينة ممثلة للمخرجات المرغوبة.

- ☒ انتقاء عناصر الاختبار بحيث تكون ممثلة لجميع عناصر المجال، ويراعى هنا تحديد أهمية كل عنصر ووزنه النسبي الذي يبنى على أساس درجة أهمية العنصر وطوله.
- ☒ صياغة بنود الاختبار بشكل علمي، وذلك على النحو التالي:

1) أن تصاغ الأسئلة بأسلوب يناسب النواتج المراد قياسها حتى يمكن عزل الصفة المراد إظهارها.
2) أن تصمم البنود بشكل ينسجم من الهدف الذي يسعى الاختبار إلى تحقيقه.
3) أن تراعى السلامة اللغوية لكل بند من بنود الاختبار.
4) أن يراعى الوضوح، وعدم الغموض في كل بند من بنود الاختبار.
5) وضع الحدود الزمنية الملائمة للاختبار، ويتم ذلك من خلال تطبيقه على عينة محدودة غير عينة البحث، ويشترط أن تكون من نفس المجتمع الأصل.
6) وضع تعليمات الاختبار بشكل يوضح للمفحوصين المطلوب منهم من حيث طريقة تسجيل الإجابات، والوقت المحدد للإجابة، مع إعطاء أمثلة لكل نوع من أنواع الأسئلة.

عناصر عملية القياس:

تشير عملية القياس إلى وجود ثلاثة عناصر أساسية متضمنة في هذه العملية، 1- موضوع القياس أو الظاهرة المراد قياسها، 2-أداة القياس التي تستخدم لإصدار حكم كمي عن الظاهرة، 3- العدد أو الرقم الذي يشير إلى درجة وجود الخاصية المدروسة بناء على قواعد مختلفة.

1- موضوع القياس (الظاهرة):

عملية القياس لا تتم في فراغ، بل تتناول ظاهرة محددة يقرر الباحث قياسها. فعندما نريد أن نجري عملية قياس، فإننا لا نتناول فرد أو نقيسه من جميع جوانب شخصيته، بل نحدد سمة أو خاصية من خصائصه ونخضعها لعملية القياس، من منطلق قابليتها لذلك، وعليه يشير سعد جلال (1985) إلى أن القياس لا يوجه إلى أشياء أو اشخاص، بل لخصائص معينة يتم تحديدها بصفة دقيقة، بحيث يمكن في النهاية تحويل الوصف الكيفي للظاهرة إلى وصف كمي يسمح بإجراء مختلف التحاليل الإحصائية لفهم الظاهرة.

ففي الميدان الفيزيائي مثلاً، لا نقيس السبورة بل نقيس طولها أو عرضها، وكذلك في العلوم السلوكية لا نقيس الفرد، بل نقيس ذكاءه، تحصيله، ميوله، أو دافعيته. فبعد أن يحدد الباحث الظاهرة، يتم ضبطها بدقة إجرائياً مع محاولة عزلها عن بقية الظواهر الأخرى خاصة مع التداخل الموجود بين الظواهر، فيصبح التحديد الواضح لمؤشرات الظاهرة الخطوة الأساسية الأولى والمهمة، والأكثر

تعقيداً في عملية القياس والتي توجه الخطوة الموالية والمتمثلة في تحديد طبيعة الأداء المناسبة لقياسها.

2- الأداة (الاختبار):

لإجراء عملية القياس لابد من توفر أداة قياس معينة تتناسب مع طبيعة الظاهرة أو السمة المراد قياسها، فالخطوة الثانية في عملية القياس بعد التحديد الدقيق الإجرائياً للظاهرة المراد قياسها هي اختيار أو بناء الأداة المناسبة. فالاختبار في علم النفس والتربية حسب صفوت فرج (1980) يعتبر بمثابة مثيرات مرتبة أعدت لتقيس بطريقة كمية في أغلب الحالات، وكيفية في بعضها لتقيس عينة من السمات أو الخصائص، والمثيرات قد تكون اسئلة شفوية، كتابية، سلسلة من الأعداد أو الأشكال الهندسية وغيرها من المثيرات التي من شانها استثارة استجابة معينة من طرف الأفراد تخضع لعملية التكميم.

والاختبار النفسي يقيس ظاهرة محددة ليعطي درجة أو قيمة كمية بمقدار الظاهرة، والباحث إما أنه يجد الاختبار جاهزا فيقرر مدى صلاحيته لقياس الظاهرة عند أفراد يمتازون بخصائص معينة، فيطبق الاختبار عليهم، ولكن قد لا يجد الباحث اختبارا جاهزا يناسب طبيعة الظاهرة أو خصائص العينة، أو ان المتوفر لا يحقق اهداف البحث. وفي هذه الحالة على الباحث ان يقوم ببناء اختبار وفق منهجية وخطوات معينة بحيث تكون بنوده عينة ممثلة فعلا للمجال السلوكي المراد قياسه بطريقة غير مباشرة. فالاختبارات في علم النفس والتربية اصناف،

فهناك اختبارات ادائية وأخرى تحريرية والبعض شفوي والآخر كتابي و تطبيقها يعطينا وصفا كميا في شكل أرقام تعتبر المكون أو العنصر الثاني لعملية القياس.

3- العدد (الرقم):

يعتبر العدد أو الرقم بمثابة اللغة المشتركة بين مختلف العلوم، ومع التطورات التي عرفتها مختلف العلوم إلا بعد أن أصبحت تتخذ من الرقم وسيلة للتعبير عن الظواهر التي تدرسها والقياس النفسي في أصله يقوم على التعامل مع السمات موضوع القياس بلغة الأرقام، فالرقم هو التعبير الكمي الدال على هذه السمات، وهو يوفر قدرا من الدقة لا يوفره الحكم الكيفي حسب صلاح أحمد مراد (2000)، حيث يمكن من خلال خصائص العدد مقارنة قدرات الأفراد بعضها ببعض ومعرفة درجة توفر خاصية ما لدى فرد أو عدد من الأفراد مما يسمح لنا باتخاذ قرارات مختلفة لا يمكن الوصول اليها بغير هذا الأسلوب، كما يصبح من الصعب تقديم تفسير موضوعي لظاهرة معينة واعطائها دلالتها دون توفر هذا النوع من المعطيات.

تصميم الاختبار:

إن الباحث في مجال علوم التربية والنفس والاجتماع يطبق مقياسا ما على عدد من الأشخاص لإعطاء كل منهم درجة تدل على مدى تحصيله أو مدى اتصافه بصفة نفسية خاصة، أو درجة اقتناعه برأي اجتماعي معين أو درجة تعصبه لجهة من الجهات. ويتضح مما سبق أن للمقاييس في حقل التربية وعلم النفس دورا

هاما كأدوات بحثية يعتمد عليها في عدة مجالات كالتقويم، والاختيـار المهنـي والتعليمـي، والتوجيه والتصنيف، والتشخيص الإكلينيكي، وفحص الاتجاهات والرأي العام...الـخ، فضـلا عن استخدامها بشكل واسع في التحقق من الفروض العلمية.

وواقع الأمر قد يصطدم الأخصائي النفسي أو الاجتماعي أو الباحث بوجه عـام في كثـير مـن الأحيان بعدم وجود مقياس مناسب للصفة أو السمة أو الخاصية المراد قياسها، أو حتـى لا يناسب أفراد عينته، ومن ثم يصبح لزاما عليه أن يقوم بتصميم مقياس يقيس تلك السـمة أو الصفة أو الخاصية ويناسب أفراد عينته.

وتعتمد عملية تصميم المقاييس في المقام الأول على القيام بعدة خطوات متسلسلة تـؤدي في النهاية إلى تجنب كثير من الأخطاء، وتتيح إمكانية إعداد مقيـاس جيـد يعتمـد عليـه في المجال المعني، وهي تحتاج تدريبا خاصا نظرا لما تستوجبه من توفر أساس نظـري وعمـلي يعين على القيام بها على الوجه الأمثل.(صديق، 2005: 82)

ويلجأ الباحث إلى تصميم وإعداد اختبار عندما يتضح لـه أن الاختبـارات المتـوفرة لا تفـي بغرضه، وفيما يلي

خطوات إعداد أو تصميم الاختبار:

تقسم خطة تصميم الاختبار إلى عدد من الخطوات المحـددة، وتتضـمن في كـل الخطـوات من البداية التي تبدأ بها وتنتهي إليها كالآتي:

1-تحديد الهدف أو الأهداف الرئيسة من الاختبار.

2- تقرير محك أو معيار الدرجة.

3- ترجمة المفاهيم أو الأهداف إلى خصائص محددة.

4- تصميم فقرات مناسبة تعبر عن هذه الخصائص.

5- اختبار مستوى الصعوبة والصدق للفقرات.

6- فعالية المشتتات (المموهات).

7- إعداد الاختبار للاستخدام.

8- تقنين الاختبار.

وفيما يلي توضيح لتلك الخطوات:

1- تحديد الهدف أو الأهداف الرئيسة من الاختبار:

إن تحديد الهدف من الاختبار من حيث ميادين استخدامه والمجتمع المراد تطبيقه عليه، على جانب كبير من الأهمية. فإن كان الهدف من استخدام الاختبار في الحصول على بيانات دقيقة، كان لا بد من استخدام الطرق الإحصائية والفنية التي تحقق هذا الغرض، إذن رغم أننا قد نبدأ من مفهوم علمي مقبول يمكن اختباره بواسطة اختبار معين، إلا إن تحديد الاختبار يتطلب مزيداً من التفصيل. فقد يكون الهدف من اختبار ما، تحديد مستويات الأفراد وفقاً لمقدار امتلاكهم

للخاصية، أو قد يكون الهدف تحديد الصعوبات أو الاضطرابات التي يعاني منها الفرد، وبالتالي تصبح مهمة الاختبار تحليل هذه الصعوبات بشكل تفصيلي وتصنيفي.

ويختلف كل هدف من هذه الأهداف في طبيعته، فالأساليب التي تهدف إلى تحديد مستوى الأفراد يكثر استخدامها في المجالات التحصيلية أو في قياس القدرات المعرفية. مما يتطلب تصنيف وتحليل العمليات المختلفة في المادة المحصلة، حيث يمكن استخدام هذا التصنيف كدليل وصفي لناتج العملية التعليمية ومدى تحقق أهدافها.

2- تقرير محك أو معيار الدرجة:

تفرض أهداف الاختبار نوع الإطار المرجعي الذي تفسر في ضوئه درجة الفرد وما إذا كان يعد جيد الأداء أم لا ويمكن التمييز هنا بين نوعين من المحكات:

أ- المحك المرجعي:

ويستخدم هذا المحك أساساً في اختبارات التحصيل فقياس التحصيل يتضمن الإشارة إلى حدود مقبولة للأداء على متصل يبدأ قطبه الأدنى بلا كفاءة على الإطلاق، وينتهي قطبه الأعلى بأداء محكم تماماً وفق تعريفات وحدات الاختبار مسبقة التحديد.

وحتى يمكن تصميم اختبار جديد ذي محك مرجعي فمن الضروري العمل وفق خطوات محددة من ذلك ضرورة وضع أشكال الأداء المطلوبة بصورة صريحة

ومباشرة. وضرورة تحديد محكات قبول الأداء وفقاً لبرنامج تحصيل الموضوع المعين، مثلاً ويجب تحليل الأداء المطلوب لعناصره المختلفة حتى يمكن تقدير جوانب تحصيل المادة أو الخبرة موضوع الاختبار، وأخيراً فإن النجاح على الاختبار يتقرر من خلال مقارنة الأداء لا بأداء آخرين، ولكن بمقارنته بحجم المطالب المحددة في الاختبار، وما أنجز منها، وسيتم لاحقاً تحديد الخطوات التفصيلية لبناء الاختبارات المحكية المرجع.

ب- المحك المعياري:

تختلف الاعتبارات ذات المحك المعياري عن الاختبارات المحكية المرجع في أن الدرجة الخاصة بالفرد لا تقارن وفق حجم سابق التحديد للمطلب موضوع الاختبار، وإنما تقارن بالأداء الخاص ببقية الأفراد، فنحن هنا نهدف للتعرف على الوضع النسبي للفرد بين مجموعة الأفراد، وبهذا تكون الدرجة الخاصة بالفرد ذات قيمة من حيث أنها تفيدنا بمدى جودة هذا الأداء بالنسبة للمجموعة المختبرة، ولكنها ليست ذات فائدة من حيث الإشارة إلى مدى جودة الأداء بالنسبة للمطالب الأساسية للاختبار أو المجال موضوع القياس. ويترتب على تقرير هذا المحك المعياري أن نتجه في تصميمنا للاختبارات إلى أن تكون الفقرات ذات قدرة على تفسير مدى الدرجات التي يمكن أن نحصل عليها من أية مجموعة من الأفراد.

ومعنى هذا أن الدرجة على المقياس في حد ذاته لا معنى لها على الإطلاق إلا إذا نسبت إلى إحدى المحكين: المحك المرجعي بأن تنسب إلى الأداء نفسه بوصفه

مرجعاً للدرجة، أو المحك المعياري بأن تنسب إلى درجات بقية الأفراد بوصفهم معياراً لهذه الدرجة.

3- ترجمة المفاهيم والأهداف إلى خصائص محددة:

المفاهيم العلمية كالذكاء والانبساط والتطرف...، تجريدات لخصائص مفترضة في الأشياء، ونستدل على هذه الخصائص من وقائع سلوكية محددة، سواء في شكل أفكار معبر عنها، أو حلول لمشكلات أو استجابة لمنبهات محددة، أو أساليب تعامل مع البيئة، أو صفات شخصية تتسم بالاستقرار.

ونقوم بترجمة هذه المفاهيم إلى خصائص محددة بصورة تسمح بصياغتها في وحدات معيارية للقياس، ففي مفهوم مثل الطلاقة يمكن ترجمته إلى أعمال محددة إذا وضع له التعريف الإجرائي المناسب، مثل (الطلاقة هي القدرة على إنتاج أكبر عدد من الأفكار ذات الدلالة استجابة لمنبه معين).

إذ يتيح هذا التعريف الإجرائي وضع المنبهات لاستدعاء استجابات معينة هي التي يقاس بواسطة كمها واستيفائها لمحك الدلالة لطلاقة المفحوص، ولم يكن متاحاً بالمثل قياس مفهوم الذكاء دون توفر التعريف الإجرائي الذي وضعه "بينيه" وترجمه إلى أعمال محددة يقوم بها الطفل. فالمطلوب في هذه الخطوة هو أن نحول الأهداف إلى سلسلة من الخطوات والأعمال، لنتمكن من كتابة الفقرات المناسبة التي يمكن أن يقدم تقديراً صادقاً للجوانب التي تقوم بقياسها، ويتم ذلك بتحديد عينة مقننة من هذا السلوك، والذي يتعين أن يمثل بشكل جيد المفهوم الذي نصمم من أجله المقياس من خلال خطوات محددة، وهي:

أ- وصف السلوك المطلوب قياسه:

فإذا كان المفهوم هنا هو ((القدرة الحسابية)) فإن وصف السلوك يكون كـالآتي: يتضـمن السلوك المطلوب قياسه القيام بحل عدد من المشـكلات الحسـابية البسـيطة دون أخطـاء، ويمكن إجراء العمليـات الحسـابية جميعهـا إمـا داخليـاً ((عقليـاً)) أو باسـتخدام الورقـة والقلـم، ويقصـد (بالعمليـات الحسـابية البسـيطة) الجمـع والطـرح والضرب والقسـمة والنسب المئوية والكسور، ولأهمية السرعة في المطلوب، والإجابة الصحيحة هـي المطلوبـة، ولذلك يمنح المفحوص وقتاً متسعاً للإجابة.

ب- تحليل السلوك المطلوب قياسه:

ويقصد به تصنيف هذا السلوك للتعرف على مكانه بين أشـكال السـلوك الأخـرى والإشـارة إلى طبيعته وإلى عموميته وقابليته للتنبؤ، فإذا طبقنـا مثـل هـذا التحليـل عـلى ((القـدرة الحسـابية)) فـيمكن أن نصـل إلى أن القيـام بعمليـات حسـابية بسـيطة يتـأثر بكـل مـن الاستعداد والتدريب أو العادات المستقرة لفترة طويلة والأفراد البارزين في هـذا المجـال لا يحتاجون لبذل مجهود كبير لتجنب الأخطاء، وحسـاباتهم دائمـاً صـحيحة بصـورة طبيعيـة، والأفراد الذين يفتقدون هذه الخاصية يقعون بالخطأ بسهولة وإجاباتهم الصـحيحة تعـود في حقيقة الأمر إلى مراجعتهم الدقيقـة وإعـادة المراجعـة، والواقـع أن الأشـخاص أصـحاب الدرجة المرتفعة يتمتعون بما يمكن اعتباره القدرة على المراجعة الآلية.

ج- تقرير خصائص الفقرات:

ويقصد بها التقدم نحو وصف نوع الفقرات التي يمكن أن تؤدي إلى صدق قياس وتقويم السلوك، فإذا تقدمنا من المثال الخاص بالقدرة الحسابية، سنجد أن قياس السلوك الحسابي يتطلب أن يتضمن الاختبار عدداً كبيراً من المشكلات البسيطة نسبياً والتي تتضمن عمليات جمع وطرح وضرب وقسمة وكسور عشرية ونسب، ويتعين أن لا يتضمن الفقرات مشكلات صعبة أو بها خدعه يجب معرفتها للوصول إلى الحل، ويجب أن تكون التعليمات الموجهة للمفحوص واضحة ومتضمنة، وأن يعمل المفحوص حسب سرعته الخاصة، ولكن دون إنفاق وقت كبير في المشكلة أو فقرة أكبر مما يجب.

4- تصميم مفردات مناسبة تعبر عن هذه الفقرات:

إن هناك أكثر من شكل وأسلوب يمكن أن تصاغ به فقرات الاختبار. وتتطلب كتابة فقرات مناسبة لاختبار جديد، تحليل الفقرات تحليلاً كيفياً من حيث شكل الفقرات ومن حيث مضمونها إلى فقرات جيدة تقيس بشكل صادق السلوك المعين كما يؤدي هذا التحليل الكيفي إلى توضيح أهمية مراعاة شروط كتابة الفقرات الجيدة.

كما يتطلب الأمر تحليلاً كمياً، يهدف لتقدير مستوى صعوبة الفقرات، وانتخاب الفقرات ذات الصعوبة المناسبة لأهداف الاختبار. ويوضح هذا النوع من التحليل علاقة صعوبة الفقرات، ويمكن التوصل إلى صدق وثبات مرتفعين للاختبار منذ البداية من خلال إجراءات تصميم الفقرات.

وتصنف الاختبارات من حيث بنيتها إلى اختبارات محددة البنية، أو الاختبارات الموضوعية التي تتميز بأن المفحوص لا يقدم فيها إجابة للفقرة، بل يختار إحدى الإجابات من بين عدد من الاختيارات التي يقدمها الاختبار، وتعتبر موضوعية حيث أن تصحيح الاختبار يتم بدون تدخل من جانب الفاحص. كما أن لهذه الاختبارات القائمة على أساس الاختيار أنواع نذكر منها:

أ- اختبارات الصواب والخطأ.

ب- اختبارات المطابقة.

ج- اختبارات الاختيار من متعدد.

وهناك نوع آخر من الاختبارات والتي تتضمن فقرات غامضة أو لا تستدعي استجابات موضوعية التصحيح بل تحتمل تأويلات وتفسيرات مختلفة قد تختلف من باحث لآخر أو وفق مبدأ للتصحيح أو آخر، وهذه الاختبارات تسمى الاختبارات الإسقاطية.

وفي فئة ثالثة نجد الاختبارات الإنتاجية والتي تتطلب من المفحوص تقديم استجابات قابلة لتصنيفات واضحة المعالم محددة بشكل مسبق مثلما نجد في الاختبارات الإبداعية أو بعض المقاييس الفردية للذكاء كاختبار الفهم أو المعلومات.

وقبل كتابة فقرات الاختبار على واضع الاختبارات أن يفكر في طول الاختبار وعدد فقراته، وأن أهداف الاختبار هي التي تحدد طوله، فإن كان لأجل مقارنة

سريعة فلا داعي للاختبار الطويل. أما إذا كان لأجل جمع بيانات دقيقة، الواجب أن يكون الاختبار طويلاً ومتعدد الأسئلة على أن يتضمن جميع مكونات الظاهرة التي يقيسها.

القواعد العامة في كتابة الفقرات للوصول إلى مستوى صدق جيد للاختبار:

- ✓ احتواء الفقرة على فكرة واحدة فقط والابتعاد عن الأسئلة المزدوجة.
- ✓ الابتعاد عن استخدام الكلمات الغريبة أو الكلمات التي تحمل أكثر من معنى واحد.
- ✓ محاولة استخدام الصيغة الإيجابية والسلبية للفقرات كلما أمكن ذلك.
- ✓ تجنب استخدام الإطلاق في العبارات مثل دائماً، أبداً، حتماً، وفي كل مكان.
- ✓ البساطة والسهولة وعدم الغموض.
- ✓ الموضوعية وإمكانية تبويب الإجابات.
- ✓ تحاشي الأسئلة الإيحائية.
- ✓ تحاشي الأسئلة التطفلية والتي تسبب إحراجاً.

5) اختبار مستوى الصدق والصعوبة لفقرات الاختبار.

هناك اعتباران أوليان يبرزان عند اختيار فقرات اختبار ما، وهي على النحو التالي:

الأول هو: اعتبار الفقرة صادقة في قياس السمة التي نرغب في قياسها، بمعنى هل تمكننا هذه الفقرة من التمييز بين الأشخاص أصحاب القدر الضئيل منها.

الاعتبار الثاني: اعتبار مستوى صعوبة الفقرة مناسبة للمجموعة من الأفراد الذين سيختبرون بهذا الاختبار.

صدق الاختبار (الفقرات):

ويعني مؤشر الصدق حسن قياس الفقرة أو تمييزها في اتساق مع بقية الاختبار أو حسن تنبؤه بالمحك الخارجي. ويمكن أن نلاحظ عدداً من المؤشرات الشائعة للصدق في الاستخدامات المختلفة من ذلك: نسبة المفحوصين الذين يجيبون على الفقرة إجابة مقبولة (تمييزية)، أو الارتباط بين الفقرة ومحك خارجي أو الارتباط بين الفقرة وبقية الاختبار وهو محك داخلي.

ونادراً ما يلجأ مصمم الاختبار إلى حساب الارتباط بين الفقرة والمحك الخارجي، وقليلاً ما يحسب الارتباط بين الفقرة وبقية الاختبار والأسلوب الأكثر شيوعاً هو حساب نسبة المفحوصين الذين يتجاوزونه ودون إجراء مثل هذا التحليل لصدق الفقرات. ومهما كانت خبرة مصمم الاختبار وموضوعيته فسيظل الاختبار في حاجة إلى فحصه لتقرير هذه الخصائص التمييزية للفقرات.

وتستخدم المجموعات المتطرفة أو المتعارضة عادة في حساب الصدق التمييزي للتفرقة بين الأسوياء وغير الأسوياء وكذلك للتمييز بين مجموعتين عمريتين مختلفتين، وذلك بأن ننتخب مثلاً الفقرات التي أجاب عليها أعلى (10%) من

المجموعة (أ) وأقل (10%)من المجموعة (ب) المتطرفة معها في السمة المقاسة وبقدر تعارض المجموعتين بقدر حدة التمييز بين الذي نتوصل إليه.

إلا إن استخدام المجموعات الشديدة التطرف بهذه الصورة يؤدي إلى خفض الثبات نتيجة للعدد الصغير من الحالات المستخدمة حيث تكون المجموعتان أ، ب عبارة عن الحالات المتطرفة في عينة واحدة مستخدمة لحساب صدق الفقرات.

ويقترح "كيلي" استخدام أعلى وأدنى (27%) من التوزيع باعتبارهما المجموعتين المتطرفتين بشرط إعتدالية التوزيع. بينما كيرتون يعتبر أن النسبة مقبولة إذا كانت تتراوح بين(27%- 33%) وإذا كان المنحنى أكثر استواءً من المنحنى الاعتدالي التقليدي. وتضيف "أنستازي" إلى أن خطأ العينة يصبح كبيراً في حالة العينات الصغيرة. ولهذا يصبح من الأفضل أن لا نحدد هذه النسب بشكل حاسم، وتعتبر النسب التي تتراوح بين (25%-33%) وافية بالغرض بصورة مقبولة، ومؤشر التمييز المقبول بصفة عامة باعتباره مؤشراً لصدق الفقرات يتم حساب معامل التمييز بهذه الطريقة بطرح عدد الأفراد الذين اختاروا الخيار الصحيح للسؤال في المجموعة الدنيا من عدد الأفراد الذين اختاروا الخيار الصحيح للسؤال من المجموعة العليا، ثم قسمة الناتج على عدد الأفراد في إحدى المجموعتين فقط.

صعوبة الاختبار (الفقرات):

إن فعالية الاختبار تتأثر بشكل مباشر في عدد من المواقف بتشتت قيم صعوبة الفقرات، وبصيغة عامة تحدد صعوبة الفقرات بحساب نسبة الأفراد الذين يجيبون

على الفقرة إجابة صحيحة، فإذا كانت الإجابة على الفقرة إما صواب أو خطأ فالأسلوب المباشر لتحديد صعوبته هو حساب متوسط الدرجة عليه أو نسبة الإجابة بصواب للمجموع الكلي للأفراد، وهي تمثل الاحتمالية العلمية لقدرة الجمهور المعني بالفقرة أو التمييز بالقدرة على الإجابة عليه إجابة صحيحة.

ولأن إمكانية توحيد مستوى الصعوبة بشكل حاسم في كل فقرات الاختبار تؤدي إلى آثار جانبية على كل من صدق الاختبار وتجانسه الداخلي، فيتعين الاهتمام في هذه الحالة بترتيب الفقرات من حيث مستوى صعوبتها من الأسهل إلى الأصعب على أن يكون مدى الصعوبة بين أسهل وأصعب فقرة محدوداً بقدر الإمكان ويدور حول نسبة 50% صعوبة ويؤدي ترتيب الفقرات من حيث الصعوبة إلى عدد من المزايا:-

من ذلك تنمية ثقة المفحوص بإجابته على الفقرات السهلة أولاً قبل أن يلتقي بالفقرات الصعبة قرب الحد المقابل لأقصى قدرته، كما يؤدي هذا الترتيب إلى خفض إمكانية تضييع جزء من الوقت في فقرات تفوق فدرة الفرد وإهماله لفقرات سهلة وفي متناوله.

ويقتصر تحليل مستوى الصعوبة على اختبارات الاستعدادات والقدرات والتحصيل، وهو أكبر أهمية بالنسبة لاختبارات الاستعدادات والقدرات، أما اختبارات التحصيل فيحدث أحياناً أن يكون تقدير خبير في الموضوع أكثر أهمية من حساب مستوى الصعوبة إحصائياً. كما أن اختبارات السرعة بكل أشكالها تخرج عن إطار مشكلة الصعوبة، إذ تتضمن أساساً فقرات سهلة، ويفضل اختيار

الفقرات ذات الصعوبة المتوسطة، وبمدى ضيق حول متوسط الصعوبة (0.50) لأنه يعطي أعلى قيمة تباين للفقرة، وهذا ما يؤدي إلى ارتفاع قيمة تباين الاختبار ككل وبالتالي ارتفاع قيمة ثباته، أما في اختبارات القوة فيفضل تدرج فقرات الصعوبة من الأسهل للأصعب، وبالتالي احتواء الاختبار على مدى واسع من مستويات الصعوبة. ويمكن الوصول إلى مقياس أفضل لمستوى الصعوبة بواسطة تحويل نسب الصعوبة إلى درجات معيارية باعتبار النسب من الأفراد الذين يجيبون إجابة صحيحة على الفقرة مقابلة للنسب تحت المنحنى الاعتدالي.

6) فعالية المشتتات (المموهات): المشتتات عبارة عن الخيارات الأخرى غير الصحيحة للسؤال ذي الاختيار من متعدد، ويفترض في المشتتات التالي:

أ- أن تكون جذابة وبالذات للأفراد الذين لا يعرفون الاختيار الصحيح، أي يتم اختيار أي مشتت من قبل مفحوص أو أكثر بنسبة لا تقل عن (5%) من المفحوصين.

ب-المشتتات عبارة عن إجابة خاطئة، لذلك فالمشتت الجيد هو الذي يكون عدد المفحوصين الذين يختارونه من الفئة العليا أقل من عددهم في الفئة الدنيا أي أنه (يميز باتجاه معاكس لتمييز الفقرة)، وكلما كانت قيمة المموه بالسالب فان هذا دليل على أن المشتت جيد.

7) إعداد الاختبار للاستخدام: بعد أن ينتهي الباحث من جميع المراحل الأساسية لتصميم الاختبار يصبح من الضروري أن يضع الاختبار في صورته النهائية التي

سيقدم بها إلى المفحوصين، من ذلك ما إذا كانت فقراته ستقدم في شكل بطاقات مستقلة تتضمن كل بطاقة فقرة واحدة أو ستوضع الفقرات في كتيب مستقل على أن يرفق بها صحيفة منفصلة للإجابة. كما يجب وضع تعليمات مناسبة للاختبار قبل تجربة الاختبار.

وفيما يلي بعض الأمور الواجب اتباعها عند وضع التعليمات للاختبار وتجربته:

أ- يجب شرح فكرة الاختبار شرحاً دقيقاً حتى لا يكون عدم فهم الأفراد لفكرة الاختبار سبباً في الإخفاق فيه.

ب- يجب أن تبرز التعليمات المهمة، وأن تكتب بخط واضح ومن المستحسن أن تكون حروفها أكبر من بقية التعليمات العادية.

ج- من الأفضل أن تحتوي التعليمات على مثال أو نوع من التمرين للمختبر قبل كل اختبار.

د- يحسن أن تكون تعليمات الأجزاء المختلفة للاختبار، متسقة أو موحدة أن أمكن. فقد ثبت بالتجربة أن توحيد التعليمات يؤدي إلى دقة أكبر وإلى ألفة المختبر بالتعليمات وتمسكه بها بينما اختلافها قد يؤدي إلى بعض الأخطاء أو إلى التضحية بالدقة.

ه- بالإضافة إلى أن التعليمات يجب أن تحتوي على طريقة الإجابة ومكانها وأسلوبها، ويجب أن تحتوي أيضاً على الزمن المطلوب مراعاته عند الإجابة إذا كان الاختبار يتطلب ذلك.

و- من الأفضل إجراء تجربة استطلاعية على عينة صغيرة لغرض التعرف على مدى الاختبار، ومدى فهم العينة لهذه التعليمات ويقصد بمدى ملائمة الاختبار أي مدى وضوح العبارات والمدة اللازمة للإجابة على الاختبار ومن ثم تعديل فقرات الاختبار في ضوء التجربة الاستطلاعية.

ومن الأمور المهمة التي ينبغي أن يراعيها واضع الاختبار هي وضع مفتاح للتصحيح، إذ أن مفاتيح التصحيح تختلف باختلاف طريقة الإجابة، أي قد تكون الإجابة على كراس الاختبار نفسه، أو إذا كانت الإجابة على ورقة إجابة خاصة ودون كتابة شيء على كراس الاختبار. وطريقة الإجابة على الورقة أفضل بكثير وأكثر اقتصادية من الإجابة على كراس الاختبار نفسه، إلا في الحالات الإستثنائية مثل اختبارات الدقة والتي تدخل فيها أرقام كثيرة أو رموز أو رسوم بيانية، أو في حالة كون التلاميذ صغار العمر وخبرتهم محدودة في الاختبارات الموضوعية.

وبعد أن تمت كتابة المفردات وأعدت التعليمات يتم تجربة الاختبار تجربة مبدئية على عينة ممثلة للمجتمع المراد عمل الاختبار له، وتسجل الملاحظات المختلفة فيما يتعلق بصلاحية التعليمات، وفهم المختبرين لها، وكذلك وقت الاختبار وتجمع كافة الملاحظات عن المفردات المختلفة، وذلك تمهيداً لتعديل ما يحتاج منها إلى تعديل.

وإثر الإنتهاء من التجارب الاستطلاعية الأولية يعمد المصمم إلى اختيار عينة أكبر من الأفراد المستهدفين في البحث بعد إجراء التعديلات المناسبة للاختبار،

ثم يجري مصمم الاختبار تجربة أساسية للاختبار يطبق فيها الاختبار على عينة من الأفراد تحمل خصائص المجتمع المراد إعداد الاختبار له، ومن خلال هذا التطبيق يتم بموجبه اختيار الفقرات الصالحة للصيغة النهائية للاختبار.

8) تقنين الاختبار:

يمكن القيام بعملية تقنين الاختبار بإتباع الخطوات الآتية:

أ. تحديد المجتمع الذي سيقنن عليه الاختبار تحديداً إجرائياً دقيقاً:

حيث أن هذه الخطوة هي الأساس لضمان صحة الخطوات اللاحقة في عملية التقنين، وتتضمن هذه الخطوة تحديد أهم خصائص وسمات المجتمع الديموغرافية، من حيث الخصائص الجغرافية والسكانية والاقتصادية، وتوزيع الفئات العمرية فيه، ونوعية التعليم، والتركيبة الاجتماعية، حيث يتم من خلال هذه المعلومات تحديد العينة التي تمثل المجتمع تمثيلاً جيداً، كما ان هذه المعلومات تمثل خصائص المجتمع والتي على ضوئها يتم تعميم نتائج الاختبار على المجتمعات الأخرى.

ب. اختيار العينة الممثلة للمجتمع وتحديد أسلوب اختيارها:

وتعتمد هذه الخطوة على الخطوة السابقة، حيث أن تحديد حجم العينة وأسلوب اختيارها يعتمد بشكل كبير على المعلومات المتوفرة عن مجتمع الدراسة كما تعتمد على الإمكانيات المادية والبشرية المتاحة، وبصفة عامة كلما كان حجم العينة كبيراً كان أفضل وأقرب إلى التمثيل الجيد للمجتمع.

ج. التخطيط الجيد والمسبق لتطبيق الاختبار: وتتمثل هذه الخطوة في وضع خطة شاملة لتطبيق الاختبار تتضمن تحديد الإجراءات والخطوات التي سوف تتبع، وتجهيز جميع أدوات ومستلزمات الاختبار، مع وضع قائمة بأسماء الأماكن التي سوف يتم تطبيق الاختبار فيها (أسماء المدارس مثلاً) وضع برنامج زمني للتنفيذ.

د. تطبيق الاختبار:

وتستلزم هذه الخطوة توحيد ظروف إجراء، وتطبيق الاختبار لجميع أفراد العينة وذلك لتحقيق مبدأ تكافؤ الفرص أمام الجميع لضمان أن الفروق التي رصدها الاختبار تعود للفروق في أداء الأفراد فقط.

ه. تحليل فقرات الاختبار:

وذلك للتعرف على مدى فعالية فقرات الاختبار، ومدى إسهامها في الحصول على خصائص سيكومترية (الصدق والثبات) تتفق مع خصائص الاختبار الجيد، وتشمل هذه الخطوة التعرف على صعوبة كل فقرة من الفقرات، وقدرتها التمييزية، كما تشمل التعرف على مدى فعالية المشتتات لكل فقرة.

و. إيجاد الخصائص السيكومترية للاختبار (الصدق والثبات):

وذلك للتأكد من أن الاختبار فعلاً يقيس ما وضع لقياسه فقط، وأن الفرق بين أداء الأفراد في الاختبار ناتج عن الأداء الحقيقي للأفراد، وليس بسبب أخطاء القياس العشوائية.

ز. إيجاد معايير الأداء:

وذلك من خلال أداء مجموعة التقنين، بغرض توفير إطار مرجعي يمكن من خلاله الحكم على أداء الفرد في الاختبار.

الشروط العلمية لانتقاء الاختبارات والمقاييس النفسية والتربوية:

إن إجراء الاختبارات يعد نوعاً من التواصل والتفاعل الاجتماعي بين طرفين أو أكثر، وهذا يتطلب اتجاها أخلاقياً مسؤولاً من جانب الفاحص، والرغبة في التعاون من جانب المفحوص، فجميع أشكال التفاعل الاجتماعي المثمر تتطلب الثقة والاحترام المتبادل بين الأطراف المتفاعلة. فإذا كان المفحوص واثقاً من النفع الذى سيعود عليه نتيجة للوقت والجهد الذى يبذله في أداء الاختبارات التي يجريها عليه فاحص متخصص، فإنه سوف يحاول جاهداً أن يتفاعل معه تفاعلاً إيجابياً، ويتبع تعليماته، لذلك تطبق على المفحوص بهدف مساعدته على فهم نفسه وصنع قراراته.

فكثير من الممارسين النفسيين يعتمدون في عملهم على عدد محدود من الاختبارات دون سواها، إما لأنها مألوفة لديهم واعتادوا على استخدامها، أو لأنها من إعداد أشخاص من ذوى مكانة علمية متميزة. ولكن هذه الأسباب لا يجب أن تكون مانعاً لهم من البحث عن اختبارات ومقاييس جديدة متنوعة. ويمكنهم تتبع هذه الأدوات الجديدة في المطبوعات التي تصدرها دور النشر ومؤسساتها المختلفة، وفي الدوريات العلمية المتخصصة.

غير أن هذا يتطلب توفر مهارات معينة لـدى الأخصـائيين النفسـيين تسـاعدهم في الحكـم على هذه الأدوات، ومدى ملاءمتها للغرض الذى سوف تسـتخدم مـن أجلـه. كـما يتطلـب توفير سبل توثيق الروابط فيما بينهم لتبادل المعلومات والخبرات في هذا المجال وغيره مـن المجالات المتعلقة بشؤون مهنتهم.

وفيما يلى بعض الأسس العلميـة والاعتبـارات الفنيـة التـي يجـب مراعاتهـا في انتقاء الاختبارات والمقـاييس التـي تسـتخدم في المجـالات المختلفـة للخـدمات النفسية:

1- توفر خصائص الاختبار أو المقياس الجيد:

فعلى الرغم من تنوع أدوات القياس إلا أنه يجب أن يتوفر فيها خصائص معينـة يصـعب بدونها الاعتماد عليها في اتخاذ قرارات صائبة تتعلق بـالأفراد. ومـن أهـم هـذه الخصـائص الموضـوعية والثبـات والصـدق ووجـود معـايير مسـتمدة مـن البيئـة التـي يسـتخدم فيهـا الاختبـار. ويمكـن أن يحصـل مسـتخدم الاختبـار أو المقيـاس عـلى بيانـات تتعلـق بهـذه الخصائص من دليل الاختبار. وهذا يجعلنا نؤكد أهمية تضـمين بيانـات وافيـة ومعلومـات كافية تتعلق بهذه الخصائص في أدلة الاختبارات، وأن يكون مستخدم الاختبار عـلى درجـة عالية من الكفاية في تقييم هذه الخصائص ومعرفـة دلالتهـا لـكي يـتمكن مـن انتقـاء أداة القياس الصالحة للغرض المعين.

ولا يجوز انتقاء اختبار أو مقياس مـن مجـرد النظـر إلى عنوانـه أو اسـمه، بـل لابـد مـن فحــص محتــوى الاختبــار فحصــا مســتنيراً والإطــلاع عــلى خصائصــه الســيكومترية

وتحديد دلالتها، والتأكد من ملاءمة معـايير للفـرد أو العينـة المسـتهدفة، والتحقـق مـن مدى صلاحية هذه المعايير.

2- مناسبة الاختبار للمفحوص:

فانتقاء الاختبار المناسب للمفحوص يعد أمراً على جانب كبير من الأهمية، فلا يجـوز أن تطبق الاختبارات التي تصلح للراشدين على تلاميذ المدارس الابتدائية أو المتوسـطة، حيـث تكون غير مناسبة للمستوى النمائى لهم، كما أن معاييرها مستمدة من عينات تختلف عـن عينات هؤلاء التلاميذ، لذلك يجب أن تشير أدلة الاختبارات إشارة واضحة إلى طبيعة عينـة تقنين الاختبار وخصائصها، من حيث العمر والنوع والمستوى التعليمـي والاجتماعـي وغـير ذلك، لكي يسترشد بها الفاحص في انتقاء الاختبار المناسب لفرد أو مجموعة معينة، وينبغي مراعاة أن وجـود مشـكلات سـلوكية، أو عجـز بصر أو ضـعف سـمع، أو حرمـان ثقـافي، أو صعوبات في اللغة تجعل الاختبار غير مناسب للمفحوص، ويجب عـدم اسـتخدام الاختبـار في هذه الحالات.

3- مناسبة الاختبار للغرض الذي يقيسه:

فعـلى الـرغم مـن تنـوع وتبـاين اسـتخدامات الاختبـارات، والمقـاييس النفسـية والتربوية، إلا أن الهدف الرئيسى منهـا جمـع بيانـات ومعلومـات تفيـد في اتخـاذ قرارات معينة تتعلق بالأفراد أو الجماعات بأقل قدر ممكن من الخطأ والصـدفة أو المخاطرة. غير أن الاختبار أو المقياس الذى يفيد في موقف معين قد لا يفيد أو لا يناسب موقفا آخر، فاستخدام الاختصاص الـنفسي أداة قيـاس مناسـبة لموقـف

معين أو صالحة لغرض معين يساعده بالطبع في اتخاذ قرارات أفضل من تلك التي يستخدمها دون استخدام تلك الأداة.

لذلك يجب تحديد الغرض من عملية القياس أو العملية الاختبارية تحديداً واضحاً وانتقاء أفضل الأدوات التي تحقق هذا الغرض، وهذا يتطلب بالطبع توفر عدد كبير من الاختبارات والمقاييس الجيدة التي تقيس مختلف السمات الإنسانية حتى يتمكن لفاحص من انتقاء أكثرها ملاءمة للغرض من القياس، وهنا تبرز أهمية تشجيع حركات القياس النفسي والتربوي وبناء الاختبارات بأنواعها المختلفة في المنطقة العربية من خلال مؤسسات متخصصة.

الشروط والإجراءات لتطبيق الاختبار:

تتعلق شروط والإجراءات لتطبيق الاختبار ببعض المبادئ والأسس المرتبطة بعملية إجراء الاختبار، والإعداد للموقف الاختباري والفحص النفسي، كما أن هناك مبادئ أساسية تتعلق بكيفية إجراء الاختبار، والتي ينبغي الإلتزام بها والعمل على مراعاتها أثناء تطبيق الاختبار.

يحتل الاختبار النفسي مكانة هامة في الفحص النفسي، وهو يشكل أداة أساسية في ممارسة الأخصائي النفسي. ان للاختبار شروطاً وقواعد ينبغي الالتزام بها أثناء التطبيق، بغية الحفاظ على موضوعية الممارسة. ولذلك، ينبغي على الفاحص أن يتبع الشروط الموضوعية لتطبيق الاختبار، ويهدف استخدام الاختبار إلى الحصول على معلومات هامة عن شخصية المفحوص، إمكاناتها

وقدراتها ودينامياتها. وتعتبر هذه المعلومات معطيات أساسية يبني الفاحص على أساسها استنتاجاته وتشخيصه للحالة.

أ- أداء المفحوص في الاختبار:

إن مفهوم " الاختبار " قد يوحي للمفحوص بعملية معينة يجب أن يخضع لها، وهذا ما يخلق لديه الشعور بالإنصياع والخضوع تجاه سلطة عليا (الموقف الاختباري)، مما يولد لديه الشعور بعدم الثقة، والقلق. ولاشك أن قلق" الاختبار " ومستوى الدفاعية يؤثران على أداء المفحوص في الاختبار، كما أوضحت ذلك " أنستازي ". في اختبارات الشخصية، وخاصة الاختبارات الإسقاطية، تصبح نوعية العلاقة القائمة بين السيكولوجي والمفحوص ذات أهمية كبيرة، نتيجة ما يوحيه الموقف الاختباري للمفحوص من إسقاطات، وتماهيات مع السيكولوجي، وهذا ما يؤدي إلى التأهيل على نوعية الاستجابة عند المفحوص.

وعلى كل الأحوال، ينبغي على السيكولوجي الانتباه لكل هذه الامور والعوامل التي يخلقها "الموقف" بحد ذاته، وعليه أن يعمل لإبعاد السلطوي تجاه المفحوص، من خلال عملية التفاعل والثقة معه، وعليه أن يلجأ أيضاً إلى تقنيات أخرى وأساليب متنوعة تساعده على تخطي كل هذه العقبات، من أجل الوصول إلى فهم أكثر لشخصية المفحوص، وقدراته الذاتية، ولذلك عليه أن يتعاون مع فريق العمل في عملية التشخيص والتقدير النهائي. بعد دراسة المفحوص من خلال مختلف الأساليب التي استخدمها الأخصائي النفسي، تأتي المرحلة النهائية من الفحص النفسي، وهي استخلاص النتائج ووضع تقرير عن المفحوص. إنها

مرحلة تثبيت المعطيات والتقرير في صورة ديناميه تعكس شخصية المفحوص الكلية، قابلة للإستخدام بهدف التوجيه أو الإرشاد أو العلاج.

ب- الشروط العلمية لتطبيق الاختبارات وتصحيحها

بعد انتقاء أدوات القياس التي تحقق بدرجة أفضل أهداف العملية الاختبارية، وتناسب المفحوص أو المفحوصين، تجرى عملية تطبيق الأدوات.

وقد يظن البعض أن تطبيق الاختبارات عملية آليه بسيطة لا تحتاج إلى توفر مهارات معينة فيمن يقوم بها،غير أننا إذا تذكرنا أن تطبيق الاختبارات يعد عملية تفاعل اجتماعى كما سبق أن أشرنا، فإننا سرعان ما ندرك تعدد المتغيرات، والظروف النفسية والفيزيائية المتعلقة بالموقف الاختبارى التي تؤثر في أداء المفحوص في الاختبارات وفي نتائج العملية الاختبارية، وإذا لم يكن القائم بتطبيق الاختبارات شخصاً مؤهلاً أو مدرباً للقيام بهذا العمل فإنه ربما لا يلتفت إلى هذه المتغيرات.

فالقلق والتوتر، وانخفاض مستوى الدافعية، وتشتت الانتباه، والضعف الحسي، وعدم الألفة بالاختبارات النفسية، والكذب وتزييف الاستجابات، والتخمين العشوائي، والاتجاه نحو الاختبار ونحو الفاحص، واحتمال عدم فهم المفحوص للتعليمات بسبب ضعف البصر أو انخفاض المستوى اللغوي، والظروف الفيزيائية غير الملائمة، تؤثر تأثيراً ملحوظاً في أداء المختبرين. ونحن عادة نطمح في أن تطبق الاختبارات في ظروف مقننة، ولكن هذا لا يتحقق بدرجة تامة في كثير من المواقف الاختبارية.

لذلك يجب أن تشتمل كراسة تعليمات الاختبارات أو ملحقاتها على توجيهات وتعليمات واضحة ومحددة عن كيفية ضبط هذه المتغيرات، وأن يقوم الفاحص بدراستها دراسة متأنية قبل تطبيق الاختبار، وكذلك الاهتمام بتهيئة المفحوصين لأداء الاختبار من خلال مقابلات فردية أو في مجموعات صغيرة لإزالة أو تخفيف مصادر القلق والتخوف وغيرها من الانفعالات السالبة، ومحاولة رفع مستوى الدافعية بدرجة تسمح للمفحوص ببذل الجهد المطلوب لأداء الاختبار، ومن ثم ينظم الموقف الاختبارى في ضوء هذا الإعداد قبل اليوم المحدد لتطبيق الاختبار. كما يجب أن تشير كراسة تعليمات الاختبار إلى الشروط اللازم توفرها فيمن سيقوم بتطبيق الاختبار.

وهناك بعض الاعتبارات التي ينبغي على الفاحص أن يراعيها أثناء تطبيق الاختبار، وهي:

1- المعرفة التامة بطبيعة كل اختبار، ومدى ملاءمته لقياس السمة التي نريد قياسها

فهناك الكثير من الاختبارات المتشابهة التي تقيس الذكاء العام، بعضها لفظي والآخر عملي، بعضها فردي والآخر جمعي، ومنها أيضاً ما يقيس القدرة على التجريد، ومنها ما يقيس الأداء العملي... ولذلك يتعين على الأخصائي النفسي أن يكون ملماً الماماً وافياً بطبيعة الاختبارات للتأكد من صلاحية كل منها لقياس الظاهرة التي يريد قياسها، وصلاحيتها لمستوى الفرد الذي سيطبق عليه من حيث مستوى سنه ومستوى تعليمه أو ثقافته وإطاره الاجتماعي – الحضاري.

2- تختلف الاختبارات فيما بينها من حيث الشروط التي تفرضها على استجابات المفحوص

فهناك اختبارات لا تسمح إلا بإجابات معينة: نعم، لا (الاستبيان)، وأخرى تفسح المجال للمفحوص أن يظهر الكثير مما لديه من تخيل وابتكار، مثل الاختبارات الإسقاطية (رورشاخ، وتفهم الموضوع). لذلك ينبغي على الفاحص التأني في اختيار الاختبار المناسب، ومدى توافقه مع الظاهرة التي يريد دراستها أو قياسها.نستنتج من ذلك، أنه ينبغي على الفاحص أن يلم كلياً بالاختبار وأن يعرف طبيعته، وحدوده، وأهدافه، وأوجه الصعوبة والسهولة فيه.

3- تعليمات الاختبار (مرتكزات الأخصائي في تطبيق الاختبار)

أثناء عملية إجراء الاختبار، على السيكولوجي أن يراعي عدة اعتبارات، أهمها:

أ- التقيد بالتعليمات وشروط الإجراء الموضوعة لكل اختبار. والهدف من ذلك هو توحيد ظروف الإجراء وإعطاء نفس الفرصة لجميع الأفراد، وعليه أن يستخدم أيضاً نفس التعابير المكتوبة أثناء طرح الفقرات، واتباع نفس السرعة أو النبرة الصوتية في إلقاء الأسئلة، ونفس الشروحات التي نصت عليها التعليمات، وهذا يعني أهمية التزام الفاحص بالشروط الموضوعية أثناء تطبيق الاختبار، فيجب أن تكون هذه الشروط هي نفسها لجميع الأفراد الذين يخضعون للموقف الاختباري. وكذلك على الفاحص أن يتقيد بالزمن المحدد بالنسبة

للإجابة على كل فقرة (سؤال) من فقـرات الاختبـار، أو للوقـت المحـدد للاختبـار ككل.

ب- يجب على الأخصائي أن يتأكد أن كل فرد في المجموعة التي يجري عليها الاختبـار أن يعبر عن امكاناته وقدراتـه، دون أيـة ضـغوط أو إيحـاءات. ولـذلك يجـب أن يطبق الاختبار في ظروف موحدة لكل أفراد العينة.

4- من المبادئ الأساسية لنجاح الأخصائي في تطبيق الاختبار إقامة علاقة تفاعليـة بينه وبين المفحوص: ولكي تتم هذه العلاقة التفاعلية على الأخصائي، أن ينتبه إليه ويصغي إلى أقواله أو أسئلته أو إلى حديثـه. وهـذه الوضـعية تتـيح للفـاحص أن يلاحظ سلوك المفحوص وتصرفاته في الموقف: من ردود الفعل، أو التردد والقلـق، أو الارتيـاح والاندفاعيـة، أي مختلـف انفعالاتـه، كـما أن هـذه العلاقـة تـؤدي بالمفحوص إلى الشعور بالراحة والطمأنينة، وبالتالي فإنها تخفف مـن حـدة قلقـه ومقاومته اللاواعية التي قد تشكل في الموقـف الاختبـاري، فوجـود تلـك العلاقـة تنعكس على المفحوص في نيل ثقته وتعاونه الإيجابي، مما تساعده على الإجابـات الصريحة والمعبرة عن حقيقة مشاعره وانفعالاته واتجاهاته، كما تتيح له الفرصـة لـكي يعطـي أقصى الأداء للاسـتجابة التلقائيـة، وليسـت المفتعلـة أو المحرفـة والمموهة.

5- البـدء بالأسـئلة أو الاختبـارات المألوفـة أو المهمـة كي لا تثـار انفعـالات المفحوص بشكل حاد أو تثيره وتحرك فيه مشاعر ووجدانات متناقضة، أو مآزم

نفسية معينة: فعادة ما تترك الفقرات أو الأسئلة " الحرجة " أو الشخصية إلى النهاية، بعد ما يكون المفحوص قد إطمأن إلى الموقف الاختباري ووثق بالفاحص، وترك العنان لتراعي أفكاره ومشاعره وخواطره، وكل ما يجول في باله.

6- عدم التعليق سلباً على الأجوبة الخاطئة (في حالة اختبارات الذكاء والتحصيل الفردية)، وعدم مقاطعة المفحوص أثناء استرساله في الحديث والتداعي: (في حالة الاختبارات الإسقاطية: رورشاخ وتفهم الموضوع). كذلك لا يجب توبيخ أو لوم المفحوص لعدم تعاونه. فعلى الفاحص أن يشجع المفحوص ويساعده على تفهم الموقف الاختباري، والهدف من الاختبار، بأنه وضع لمساعدته وتعزيز قدراته الخاصة، وخصوصاً عندما يكون المفحوص في حالة قلق وتوتر، أو تردد أو فاقد الثقة بالنفس، وخاصة عند الأطفال.

7- يجب أن يكون الأخصائي النفسي واعياً تماماً، أثناء الموقف الاختباري، لدوافعه واحتياجاته الشخصية، حتى لا تتدخل هذه العوامل الخاصة في الموقف تجاه المفحوص، وتنعكس سلباً على استجاباته مما تؤثر على النتائج.

8- يجب الانتباه لتأثير السيكولوجي على المفحوص وتأثير ذلك على الإستجابة والنتائج. فمظهر السيكولوجي وشخصيته، وأسلوبه ونبرات صوته، وحركاته وانفعالاته أثناء المقابلة، أو الفحص قد تؤثر على المفحوص. فالسيكولوجي قد يطابق أحياناً صورة لا شعورية لدى المفحوص (صورة

سلطوية يشعر المفحوص نحوها بالعداء والكراهية، أو صورة مثالية يشعر نحوها بـالتعلق والحب والإنجذاب)، مما يؤثر أيضاً على النتائج.

9- مراعاة الظروف المادية المحيطة بأداء الاختبار، مثل التهويـة والإضـاءة والمكـان المـريح، والتأكد من أن جميع الأفراد يسمعون تعليمات الاختبار بوضوح، وكذلك مراعـاة الظـروف النفسية والتأكـد مـن أن المفحـوص لـيس مراهقـاً أو مريضـاً متألمـاً، فهـذه الأمـور الماديـة والنفسية تؤثر على استجابة المفحوص ونتائجه.

10- يجـب أن تكـون التعلـيمات سـهلة وواضـحة ومفهومـة لـدى الجميـع، وهـي تقـدم فقرة تلو الفقرة وبهدوء، كما تتطلب معرفة الفقرة التي يجب التوقف عندها. كـل هـذه الأمور تساعد المفحوص على الاستيعاب التام للاختبـار، ومـا هـو المطلـوب منـه في عمليـة الأداء.

الفوائد العملية لتطبيق الاختبارات النفسية:

تتعدد فوائـد القيـاس الـنفسي والتربـوي بتعـدد فـروع علـم الـنفس الاساسـية والفرعيــة، وتختلــف تطبيقاتــه بــاختلاف مجــالات البحــث فيــه، كــما تعــد الاختبارات النفسية والتربوية مصدر معلومات يتخذ على أساسها مجموعة من القرارات منها:

أ- قرارات ذات طبيعة تنبؤية (من خلال استخدام نتائج الاختبـارات في التخطـيط ورســم التوقعــات المســتقبلية لســلوك محتمــل لــه علاقــة مبــاشرة بالســلوك المقاس).

ب- قرارات ذات طبيعة تشخيصية.

ج- قرارات ذات طبيعة تقويمية.

د- قرارات ذات طبيعة انتقائية.

إخراج الاختبار:

يتضمن إخراج الاختبار من قبل معد الاختبار مراعاة الآتي:

1- أن يحتوي على تعليمات الإجابة عن الاختبار.

2- أن تكون طباعة الأسئلة واضحة وخالية من الأخطاء المطبعية والعلمية والإملائية.

3- الفصل بين التعليمات والأسئلة.

4- الفصل بين كل سؤال والذي يليه بمسافة معقولة.

5- أن لا يجزأ السؤال على صفحتين متتاليتين.

6- الفصل بين كل نوع أو شكل من أشكال الأسئلة والشكل الآخر.

7- الفصل بين مقدمة السؤال وبدائله بمسافة معقولة.

تفسير الاختبار وتصحيح النتائج:

يتضمن كل اختبار معلومات وإرشادات عن كيفية التأويل، وكيفية وضع العلامات واستخراج النتائج. وهناك عدة اعتبارات هامة على السيكولوجي مراعاتها أثناء التصحيح وتفسير النتائج، من أهمها ما يلي:

1- على الأخصائي أن يتنبه لإمكانيات التحيز الواعية واللاواعية، وقد يتدخل التحيز أحياناً نتيجة ميول واتجاهات وقيم السيكولوجي تجاه طائفة معينة، أو تجاه أيديولوجية معينة، أو تجاه جنس معين أو عرق، أو منطقة معينة. وكذلك الحال بالنسبة للتوقع المسبق، إذ غالباً ما يؤدي إلى رفع الدرجة أو انخفاضها إنطلاقاً من الحكم المسبق الذي كونه المفحوص للوهلة الأولى.

2- هناك أيضاً عامل التعب والإرهاق أو التسرع في إعطاء الدرجات، ولذلك من الأفضل أن يراجع

الفاحص حساباته أكثر مرة للتأكد من صحتها.

3- على السيكولوجي أن يفسر النتائج من خلال الحالة العامة للمفحوص: الجسدية والنفسية والمزاجية، وكذلك من خلال الإطار الثقافي والحضاري الذي تم فيه العمل، نظراً لما لها إنعكاسات على موقف المفحوص وأسلوبه.

4- ينبغي على الأخصائي أن يلم بالاختبار كلياً، وأن يعرف طبيعته وحدوده، ويعمل من ثم على تفسير درجات المفحوص انطلاقاً من تلك المعرفة.

5- لا يجوز الإستناد إلى نتيجة اختبار واحد للحكم على المفحوص. بل من الأفضل اللجوء إلى أكثر من مصدر أو الرجوع إلى إجابات ودرجات على اختبارات أخرى، وتكوين حكم عام من خلال مقارنة مختلف النتائج للمفحوص.

هذه التعليمات الضرورية والأساسية والتي ينبغي على السيكولوجي مراعاتها وأن يتبعها بدقة متناهية، وليس له أن يغير أو يعدل فيها، إذ أن أي تبديل في كيفية إجراء الاختبار يؤدي إلى تغيير في النتائج، وتصبح مقارنة نتائج المفحوص بنتائج فرد آخر في نفس العينة غير دقيقة.

وعلى الرغم من التقيد بالتعليمات وإتباعها إلا أن هناك مجالاً أيضاً يستخدم فيه الأخصائي النفسي خبرته وممارسته الشخصية، إذ هناك شيء من المرونة حسب ما يقتضيه الموقف، وتظهر هذه المرونة خاصة مع الأطفال، لكسب ثقتهم وإزالة مشاعر الخوف والارتباك لديهم. فالغاية الأساسية من الاختبارات هي تجميع معلومات عن حالة الفرد، وكشف إمكانياته وقدراته، وفهم نواحي القوة والضعف عنده، بغية الوصول به إلى حياة ومواقف أكثر فعالية.

الفصل الثالث

مقدمة:

يعد الصدق من أهم خصائص الاختبارات والمقاييس التربوية والنفسية، فصدق الاختبار يتعلق بالهدف الذي يبنى الاختبار من أجله، وبالقرار الذي يتخذ استناداً الى دراجاته، فالصدق إذن يتعلق بمدى فائدة القياس في اتخاذ قرارات تتعلق بغرض أو اغراض معينة. فعلى الرغم من أن الثبات يعني اتساق الدرجات، فإن الصدق يتعلق بدرجة تحققنا من تفسير درجات الاختبار واستخدامها. ولذلك يمكن أن تتميز درجات الاختبار بالثبات أو الاتساق، ولكن لا يعني ذلك أن الاختبار يكون صادقاً.

ويقصد بثبات الاختبار إعطاء نفس النتائج إذا ما أعيد على نفس الأفراد في نفس الظروف، ويقاس هذا الثبات إحصائيا بحساب معامل الارتباط، والموضوعية في الاختبارات والمقاييس، إخراج رأي المصحح أو حكمة الشخصي من عملية التصحيح ، أو عدم توقف علامة المفحوص على من يصحح ورقته ، أو عدم اختلاف علامته باختلاف المصححين ، كما تعني أيضا أن يكون الجواب محددا سلفا بحيث لا يختلف عليه اثنان كما هو الحال في الأسئلة الموضوعية ، والموضوعية صفه أساسية من صفات الاختبار الجيد عليها يتوقف ثبات الاختبار ثم صدقه.

الخصائص السيكومترية للاختبارات: (الصدق،...الثبات،... الموضوعية).

مفهوم الصدق:

يقصد بصدق الاختبار اذا كان هناك دلائل مقبولة حول ما يقيسه وكيفية قياسه، ويختلف كل اختبار عن الاخر فيما يقيسه من خصائص نفسية، وعلى الرغم من أن كل اختبار يحمل مسمى يشير إلى ما يقيسه، إلا أن هذا المسمى غالبا ما يكون عاما وغير محدد، كما أنه لا يعد في حد ذاته دليلا على ما يقيسه الاختبار، ونحن فقط نقبله كأداة تعريفية للاختبار.

نتيجة لهذا يصبح من الضروري التحقق من صدق الاختبار بأساليب منهجية محددة، قبل الشروع في الاعتماد على النتائج المستخلصة منه سواء كان اعتمادا عليه لأغراض علمية أو تطبيقية.

ولا يعد صدق الاختبار خاصية مطلقة، بمعنى أن الاختبار إما أن يكون صادقا أو غير صادق، الصدق خاصية نسبية ويصدق الاختبار في قياس ما يقيسه بقدر معين، كما يختلف صدق الاختبار الواحد من مجتمع إلى آخر، نتيجة للفروق الحضارية التي تسبغ السلوك بدلالات مختلفة.

قد يختلف صدق الاختبار أيضا بين فترة زمنية وأخرى، فنتيجة للتقدم في التعليم والتطور الاجتماعي والثورة التكنولوجية الاتصالية، أصبحت البنود التي كانت صادقة في القياس في العشرينات وثلاثينات القرن الماضي لا تصلح الآن، ولا يمكن استمرار وجودها في مقياس صادق للذكاء في القرن الواحد والعشرين.

لكل هذا يتطلب التحقق من صدق اختبار ما قبل استخدامه فحصا متأنيا لمضمونه وللاعتبارات الحضارية التي تضفي على البنود دلالة معينة، ويبدو هذا الاعتبار على درجة كبيرة من الأهمية عند نقل أو ترجمة اختبار من حضارة ذات خصائص معينة إلى حضارتنا المحلية التي تتسم بخصائص نوعية واضحة تنعكس على القيم وأساليب السلوك السائدة ودلالتها الاجتماعية، كما يتعين أن يضع الباحث في اعتباره عند تقدير صدق اختبار ما، ان الامر يتطلب مراجعة اجراءات الصدق بين فترة وأخرى لإعادة تقدير تأثير المتغيرات الزمنية، التي تؤثر بشكل مباشر في المفاهيم التي يتضمنها الاختبار.

ولعل ابسط تعريف للصدق، هو أن يقيس الاختبار، الأداة، ما وضع لقياسه، ولا يقيس شيئا آخر. ويقصد به الدرجة التي يحقق فيها الاختبار الأهداف التي وضع من أجل قياسها: أي أن الاختبار يعتبر صادقا عندما يقيس ما ينبغي قياسه فعلاً.

وللقياس أسس ومبادئ لا بد من توافرها، من بينها أنه يجب أن يكون قياساً موضوعياً، بمعنى أن لا يتأثر بالجوانب الشخصية للفاحص كآدائه وميوله الذاتية. فالموضوعية تقتضي أن تصف قدرات الفرد كما هي موجودة فعلاً لا كما نريدها أن تكون.

ويرتبط الصدق ارتباطاً مباشراً بالثبات، ولكي يكون المقياس صادقاً يجب أن يكون ثابتاً، أي أن يعطي نفس الدرجات عند تطبيقه عدة مرات على نفس العينة من الأفراد.

ونجمع الأدلة على الصدق من عدة طرق، وتتفاوت أهمية هذه الطرق حسب طبيعة الصفة المقاسة وحسب الغرض من استخدامها، وبعض هذه الطرق كيفي وبعضها كمي، وتجمع بعض هذه الأدلة قبل تجريب المقياس والبعض الآخر يتم جمعه بعد التجريب.

من شروط الاختبارات الجيدة أن تتمتع بالثقل العلمي ونعني به:

أولا: صلاحية الاختبارات والقياسات: وذلك من خلال توفر:

1- القدرة التمييزية (التفريقية) (المقارنة الطرفية) للاختبارات.

2- معامل السهولة والصعوبة.

ثانيا: المعاملات العلمية: المتمثلة بـ:

1- الصدق

2- الثبات

3- الموضوعية

4 –المعايير إذا كانت اختبارات ذكاء مثلا

اولاً: صلاحية الاختبارات والقياسات:

1 - القدرة التمييزية (التفريقية) للاختبارات: وتعني

قدرة الاختبار على التمييز بين الأفـراد ذوي الدرجـة العاليـة في الصـفة أو الخاصـية المراد قياسها والأفراد الحاصلين عـلى درجـات منخفضـة فيهـا، والهـدف مـن هـذه الخطوة هو الإبقاء على الفقرات أو الاختبارات ذات التمييز العالي والجيـدة فقـط، ويجب أن لا يفهم هنا أن الأفراد ذوي المسـتوى الضـعيف لا يـؤدون أو لا يجيبـون على هذه الفقرة. بل أن تكون نسبة المجيبين عليهـا مـن الأقويـاء (الجيـدين) أعـلى من الضعاف (ذوي المستوى الضعيف) بصـورة واضـحة، وذلـك لان الفقـرة التـي لا يجيب عليها جميع المختبرين على اختلاف مستوياتهم لاقيمة لها لأنهـا لا تسـتطيع التمييز بينهم.

ويتطلب إيجاد معامل التمييز للمفردات الاختبارية الآتي:

- تطبيـق الاختبـار عـلى العينـة وإيجـاد الدرجـة التـي حصـل عليهـا كـل فـرد في الاختبار.

- ترتيب الدرجات من الأعلى إلى الأدنى للمجموعة (العينة) ككل.

- تحديد مجموعة الدرجات العالية عـن طريـق أخـذ مـا نسـبته (27 %) مـن مجمـوعهم الأصـلي، وكـذلك لمجموعـة الـدرجات المنخفضـة،وبذلك تشـكل المجمـوعتين مـا مجموعـه (54%) مـن المجمـوع الكـلي للمختـبرين أمـا النسـبة المتبقيـة وهـي (46%) فهـي تمثـل الدرجات المتوسطة للمختبرين.

- هـذه الخطـوة في بعـض الاختبـارات كالبدنيـة والمهاريـة نقـوم بالمقارنـة بـين المجموعتان عن طريق المتوسطات الحسابية والانحرافـات المعياريـة للاختبـارات

باستخدام اختبار (ت)، ومقارنة القيمة المحسوبة بقيمة جدولية تستخرج من جدول (ت)، فكلما كانت القيمة المحسوبة اكبر من الجدولية كانت هناك فروق معنوية تشير لقدرة الاختبار التمييزية أو التفريقية.

وفي اختبارات أخرى كالنفسية أو المعرفية نقوم بإيجاد عدد الأفراد الذين أجابوا إجابة صحيحة على الفقرة في كل من المجموعتين، ثم نقوم بحساب النسبة المئوية لدرجة تمييز الفقرة وفق المعادلة الآتية:

عدد الإجابات الصحيحة للفئة العليا-عدد الإجابات الصحيحة للفئة الدنيا÷عدد أفراد إحدى الفئتين× 100

ويرى المختصون بالقياس والتقويم الآتي:

- ✓ الفقرة التي تحصل على نسبة مئوية سالبة (-30%) هي فقرة غير مميزة.
- ✓ الفقرة التي تحصل على نسبة مئوية موجبة وأقل من 40 % هي فقرة ذات قدرة تمييزية ضعيفة.
- ✓ إذا تراوحت النسبة بين (40 % - 60 %) هي فقرة ذات قدرة تمييزية متوسطة.
- ✓ الفقرة التي تزيد نسبتها عن 60% هي فقرة جيدة التمييز.
- ✓ كلما اقتربت النسبة من 100 % فهذا يشير إلى قدرة ممتازة على التمييز.

وقد يستغني البعض عن النسبة المئوية بالآتي:

معامل التمييز = عدد الإجابات الصحيحة عن الفقرة في المجموعة العليا- عدد الإجابات الصحيحة عن الفقرة في المجموعة الدنيا ÷ عدد أفراد أحد المجموعتين.

ولتحديد إمكانية قبول أو رفض الفقرة في ضوء معامل تميزها وضع إبيل(Ebel,1979) مجموعة قواعد بعد أجراء العديد من الدراسات وهي:

- ❖ إذا كان معامل التميز أكبر من 0.40 فان الفقرة تعتبر ذات تميز عالي وممتاز.
- ❖ إذا كان معامل التميز بين (0.30 – 0.39) فان الفقرة تعتبر ذات تميز جيد.
- ❖ إذا كان معامل التميز بين (0.20 – 0.29) فان الفقرة تعتبر ذات تميز جيد إلى حد ما (فقرات حدية تحتاج إلى تحسين).
- ❖ إذا كان معامل التميز أقل من 0.19 فان الفقرة ضعيفة وينصح بحذفها.

2- معامل السهولة والصعوبة:

إن معامل السهولة يعني عدد الإجابات الصحيحة على السؤال مقسوما على مجموع الإجابات الصحيحة والخاطئة، أما معامل الصعوبة فهو عدد الإجابات الخاطئة على السؤال مقسوما على مجموع الإجابات الصحيحة والخاطئة.

ثانياً: المعاملات العلمية:

1- الصدق:

يعد الصدق من أهم شروط الاختبار الجيد وهو يعني:

- المدى الذي يحقق به الاختبار أو أي متغير آخر الغرض الذي وضع من أجله.

- ما يقيسه الاختبار وإلى أي حد ينجح في قياسه.

إن صدق الاختبار في قياس ما وضع من أجله يكون بالنسبة لناحيتين هما:

- قياس السمة المراد دراستها أو الوظيفة التي يقيسها.

- طبيعة العينة أو المجتمع المراد دراسة السمة كعينة مميزة لأفراده.

- ارتباط الاختبار ببعض المحكات.

الصدق نسبي بمعنى أن الاختبار يكون صادقا بالنسبة للمجتمع الذي قنن فيه، فاختبار الركض 1500 م قد يكون صادقا لقياس مطاولة الجهاز الدوري التنفسي لطلبة الجامعة، في حين لا يكون على نفس الدرجة من الصدق إذا استخدم نفس الاختبار لقياس نفس القدرة للمرحلة الابتدائية، وعليه فالصدق ليس أمرا مطلقا بل يختلف من اختبار لآخر، بحيث لا نستطيع أن نقول أن الاختبار صادق أو غير صادق، بل نقول انه صادق بدرجة ما.

كذلك الصدق نوعي أي أن الاختبار يكون صالحا لقياس ما وضع لقياسه دون غيره، وتختلف الاختبارات في مستويات صدقها تبعا لاقترابها أو ابتعادها من تقرير تلك الصفة التي تهدف إلى قياسها.

ويتحدد صدق الاختبار عادة بمعامل صدقه الذي يعني معامل الارتباط بين درجات الأفراد في الاختبار ودرجاتهم في المحك،أي أن معامل الصدق هو أحد تطبيقات معامل الارتباط وقيمته الحقيقية والواقعية تنحصر بين (الصفر إلى +1) أي عدم وجود ارتباط سالب.

أنواع الصدق:

1- الصدق الظاهري:

يعتبر من اقل الأنواع أهميةً واستخداماً، ويعتمد على منطقية محتويات الاختبار ومدى ارتباطها بالظاهرة المقاسة. وهو يمثل الشكل العام للاختبار أو مظهره الخارجي من حيث مفرداته وموضوعيته ووضوح تعليماته. وقد يطلق عليه اسم (صدق السطح) كونه يدل على المظهر العام للاختبار.

وهذا النوع يتطلب:

- البحث عما (يبدو) أن الاختبار يقيسه.
- الفحص المبدئي لمحتويات الاختبار.
- النظر إلى فقرات الاختبار ومعرفة ماذا يبدو أنها تقيس، ثم مطابقة ذلك بالوظائف المراد قياسها. فإذا اقترب الاثنان كان الاختبار صادقا سطحيا.

وحساب هذا النوع يتطلب التحليل المبدئي لفقرات الاختبار، ولمعرفة ما إذا كانت تتعلق بالجانب المقاس وهذا أمر يرجع إلى ذاتية الباحث وتقديره، وهنا تكمن المحاذير.

2- صدق المحتوى (المضمون)

يعتمد هذا النوع من الصدق على فحص مضمون الاختبار فحصا دقيقا، وهو يعني مدى جودة تمثيل محتوى الاختبار لفئة من المواقف أو الموضوعات التي يقيسها، فيعتبر الاختبار صادق، إذا مثلت تقسيماته وتفرعاته تمثيلا سليما. ويستخدم هذا النوع في تقويم اختبارات التحصيل والكفاية والتي تقيس مدى إتقان اللاعب لجانب معين أو مدى ما حصله من التدريب أو المعارف أو المعلومات، وللتحقق من صدق المحتوى لأي اختبار نتبع ما يلي:

- تحديد السمة أو الظاهرة أو الخاصية قيد البحث تحديدا منطقيا بالتحليل الشامل
- التعرف على أبعاد السمة أو الظاهرة أو الخاصية المقاسة، وأهمية كل جزء فيها والوزن النسبي لكل جزء أو بعد من هذه الأبعاد وذلك بالنسبة للاختبار ككل.
- وضع مفردات الاختبار بما يتفق مع الأبعاد أو الأجزاء التي استقر عليها الرأي في ضوء المرحلتين السابقتين. مثال: في اختبار لقياس القوة نقوم بتحليل القدرة العضلية لأشكالها (القوة القصوى، والانفجارية، المميزة بالسرعة، مطاولة القوة) ثم نرشح اختبارات لتغطية هذه الأبعاد في ضوء

الوزن النسبي لأهميتهـا، وتقـديرات الخـبراء أو الحكـام هنـا هـي المحكـات التـي تستخدم لتحديد الصدق.

3- الصدق المرتبط بالمحك (الصدق التجريبي)

المحك هو معيار نحكم به على اختبار أو نقومـه، وقـد يكـون مجموعـة مـن الدرجات أو التقديرات أو المقاييس صمم الاختبـار للتنبـؤ بهـا أو الارتبـاط معهـا كمقياس لصدقها. والمحك هو مقياس موضـوعي تـم التحقـق مـن صـدقه لـذلك نقارن بينه وبين المقياس الجديد للتحقق من درجة صدق ذلـك المقيـاس، وذلـك عن طريق معامل الارتباط بينهما. والصدق التجريبي يعتمد على إيجـاد معامـل الارتباط بين الاختبار الجديد واختبـار آخـر سـبق إثبـات صـدقه أو محـك وكلـما اقترب معامل الارتباط بينهما من واحد صحيح كأن يكـون 0.88 مـثلا كلـما كـان الاختبار الجديد صادقا.

ويعتبر هذا النوع من الصدق من أفضل الأنواع وأكثرها شيوعا، ويصنف وفقا للغرض من استخدامه إلى نوعين هما:- الصدق التنبؤي والصدق التلازمي: ويمكن التمييز بين هذين النوعين في ضوء الفترة الزمنية بين الاختبار والمحك، والهدف من الاختبـار هـل هو تحديد الحالة الراهنة (صدق تلازمي) أو التنبؤ بنتيجة معينة في المستقبل (صدق تنبؤي).

أ- الصدق التنبؤي: يدل هذا النوع من الصدق عـلى مـدى الصـحة التـي يمكـن أن نتوقـع بهـا خاصـية أو قـدرة معينـة لـدى الأفـراد مـن خـلال اختبـار يفـترض أن يقـيس

هذه الخاصية. يعتبر هذا النوع من الصدق مؤشرا لنتيجة معينة في المستقبل، حيث يقوم على أساس المقارنة بين درجات الأفراد في الاختبار وبين درجاتهم على محك يدل على أدائهم في المستقبل، ويعتبر الاتفاق (معامل الارتباط) بين درجات الاختبار ودرجات المحك هو معامل صدق الاختبار. وعليه فهو عبارة عن عمليات يمكن من خلالها حساب الارتباط بين درجات الاختبار وبين درجات محك خارجي مستقل. مثال اختبار القدرات للطلاب المتقدمين لكلية التربية الرياضية، والقدرة على الاستمرار بالدراسة معامل الارتباط العالي بين الاثنين مؤشر صدق تنبؤي يراعى في هذا النوع من الصدق:

- حساب القيمة التنبؤية للاختبار.
- الاعتماد على فكرة أن السلوك له صفة الثبات النسبي في المواقف المستقبلية.
- تنبؤ يحتاج إلى فترة بين تطبيق الاختبار ثم جمع البيانات عن المحك في فترة تالية للاختبار.

ب- الصدق التلازمي: يمثل الصدق التلازمي العلاقة بين الاختبار ومحك موضوعي تجمع البيانات عليه وقت أو قبل أجراء الاختبار. أي التعرف على مدى ارتباط الدرجة على الاختبار بمحكات الأداء الراهنة أو مركز الفرد حاليا. يستخدم عندما يتلازم تطبيق الاختبار وتطبيق المحك معا ويصبح الهدف هو معرفة عما إذا كان كل من الاختبارين يقيسان خصائص قائمة بالفعل في وقت واحد، وذلك بهدف تقدير الحالة الراهنة، وهو من أنسب الأساليب ملائمة

للاختبارات التشخيصية فإعداد اختبار لقياس السرعة لو ارتبط بدرجة أو تقدير المدرب أو المدرس لأفراد العينة، فان معامل الارتباط العالي مؤشر صدق تلازمي.

ومن شروط المحك الجيد:

- ✓ أن يكون متعلقا بالوظيفة التي وضع الاختبار لقياسها.
- ✓ أن المقياس كمحك يجب أن يهيئ لكل شخص نفس الفرصة لاخذ درجة عادلة (البعد عن التحيز).
- ✓ أن يتوفر في المحك خاصية الثبات.
- ✓ أن يكون المحك موضوعيا.

عيوب الصدق المرتبط بالمحك:

- انه يعتمد على صدق الميزان أو الاختبار المرجعي، فإذا كان هذا الاختبار غير صادق أو مشكوك في صدقه يؤثر بذلك على الاختبار المراد معرفة صدقه.
- صعوبة ضبط الميزان بالنسبة لإيجاد الصدق.

4- صدق التكوين الفرضي (الصدق التكويني)

يقصد بالتكوين : مفهوم جوهري في العلم، وهو يعبر عن انطلاقنا من خبرة الإحساسات الفورية والمباشرة إلى الصياغات العامة، وهي صياغات ضرورية لتشكيل قوانين العلم والتكوينات، وهي التي تسمح لنا بالتعميم من تجربة جزئية

وعيانية إلى قاعدة عامة تتضمن تنوعا من الاشياء والأحداث الحاضرة والمستقبلية.

كــل التكوينــات ذات العلاقــة بــالواقع، بالأشــياء والأحــداث الملاحظــة والخاصــة المميزة للتكوينات النفسية هي انها دائما ذات علاقة مباشرة أو غير مباشرة بالسلوك أو الخبرة.

يقدر صدق التكوين بفحص أي الخصائص يقيس الاختبار بمعنى أن نحدد المفاهيم التفسيرية والتكوينات النظرية المعينة المسؤولة عن الأداء على الاختبار، ويتطلب فحص صدق التكوين مزيجا من التناول المنطقي والتجريبي واساسا تقوم دراسات صدق التكوين بفحص النظرية القائمة خلف الاختبار، وينطوي هذا الاجراء على ثلاث خطوات:

1- يبحث الفاحص عن أي الفروض يمكن وضعها، وتقوم بتفسير الدرجة المرتفعة والدرجة المنخفضة على الاختبار.

2- يقوم بجمع بيانات واقعية لاختبار هذه الفروض.

3- يقوم في ضوء ما تم جمعه من دلائل بتقديم استدلالات معينة حول ما اذا كانت النظرية مناسبة لتفسير البيانات التي أمكن جمعها أم لا.

ومن إجراءات أو أساليب حساب صدق التكوين الفرضي:

- الفـروق الفرديـة: الأفـراد يختلفـون بمـا لـديهم مـن سـمات وصـفات و ...الـخ ويختلفون بوصفهم أفراد وأعضاء في جماعات.
- التغيير في الأداء: هو دراسة الفـروق في الأداء الخـاص بالعينـة نفسـها عـلى مـدى فترات زمنية مختلفة.
- الارتباط باختبارات أخرى: إن الصدق لا يُحدَد فقط بالارتباط مع اختبارات تقيس نفس السمة، أي عـن طريـق التشـابه. وإنمـا بالارتبـاط مـع اختبـارات لا تقـيس السمة نفسها أي مختلفة.
- الاتساق الداخلي: يؤدي هذا الأسلوب إلى الحصول على تقدير لصـدقه التكـويني، وذلك من خلال إيجاد معامل الارتباط بين نتيجة كل فقرة في الاختبار على حـدة مع نتيجة الاختبار ككل (المجموع الكـلي) وكـذلك الارتبـاط بـين كـل فقـرة مـن فقرات الاختبار والبعد الذى تنتمي إليه إذا كان الاختبار يتكون من أبعاد.

5- الصدق العاملي:

ويعتمد هذا النوع من الصدق على منهج التحليل العاملى الذى يقـوم علـى تحليـل مصـفوفة معاملات الارتباط بين الاختبارات والمحكات المختلفة من أجـل الوصول إلى العوامل التـى أدت إلـى إيجـاد هـذه المعـاملات، والتحليـل العـاملى اصطلاح يمثل عدداً كبـيراً مـن الأسـاليب الرياضـية المختلفـة لتحليـل العلاقـات البينية بين مجموعة من المتغيرات ولتفسير هذه العلاقات فى شكل عدد أقل مـن التغـيرات يطلـق عليهـا اسـم العوامـل، العامـل هـو متغـير فـرضى يـؤثر عـلى

الدرجات فى واحـد أو أكثـر مـن المتغيـرات المشاهدة ، وعندما يستخدم التحليـل العـاملى على مصفوفة ارتباطات، فإن الاختبارات التى تتـأثر بعوامـل معينة يقـال إن لهـا تشـبعات عاملية عالية أو أنها تتشبع بدرجـة عاليـة عـلى هـذه العوامـل، ولهـذه الطريقـة أهميتهـا الكبرى فى تحليل عدد كبير من الاختبارات تحليلاً علمياً دقيقـاً يـؤدى إلـــى الكشـف عـن أقوى تلك الاختبارات بالنسبة لأي ميزان.

6- الصدق الذاتي:

ويطلق عليه أيضا مؤشر الثبـات وهـو صـدق الـدرجات التجريبيـة بالنسـبة للـدرجات الحقيقية التي خلصت من شوائب أخطاء الصدفة، ومن ثم فان الـدرجات الحقيقيـة هـي الميزان أو المحك الذي ينسب إليه صدق الاختبار. وبما أن ثبات الاختبار يعتمد على ارتباط الدرجات الحقيقية للاختبار بنفسها إذا أعيد الاختبـار عـلى نفـس المجموعـة، لهـذا كانـت الصلة وثيقة بين الثبات والصدق الذاتي، فهو يحسب من جذر الثبات. ويجب ملاحظـة أن النتيجة المستخلصة تمثل الحـد الأقصى المتوقـع للصـدق، ولـيس القيمـة الحقيقيـة لصـدق الاختبار.

العوامل التي تؤثر في الصدق:

1. طول الاختبار: يزداد صدق الاختبار بزيادة مكوناته سواء عبـارات أو أسـئلة أو اختبارات...الخ
2. ثبات الاختبار: يتأثر الصدق بقيمـة الثبـات، لـذلك فالنهايـة العظمـى للصـدق لاتزيد عن الجذر التربيعي لمعامل الثبات للاختبار.

3. ثبات الميزان أو المحك: يزداد الصدق تبعا لزيادة ثبات المحك ويتأثر بالقيمة العددية للمحك.
4. التباين: يتأثر الصدق بتباين درجات الاختبار فزيادة أو نقصان الفروق الفردية تؤثر على الصدق.

زيادة درجة الصدق:

لابد أن نعلم في البدء أنه ينبغي لمصمم الاختبار أن يكون ملما بثلاثة جوانب: المعرفة بالموضوع أو التخصص الذي يكتب فيه، والإلمام بفنيات القياس والمهارة الكتابية.

وهاهنا تلخيص لما سبق وتم الإشارة إليه لأهم العوامل التي يسترشد بها لتحقيق قياس صادق. فينبغي أن تكون أسئلة الاختبار ذات علاقة بالصفة المراد قياسها وألا تقيس جوانب أخرى لا علاقة له بالصفة. ولابد أن تركز الأسئلة على الجوانب المهمة من الصفة دون السطحي والهامشي منها. وعلى الاختبار أيضا أن يكون شاملا وممثلا لكافة أبعاد الصفة ومجالاتها. أضف إلى ذلك فإن ثبات الاختبار وموضوعيته من الركائز الأساسية لصدق الاختبار فلابد من أخذ العوامل التي ترفع من مستوى الثبات بالاعتبار، ومن تلك العوامل:

1. وضوح الأسئلة
2. تقليل فرص التخمين
3. زيادة عدد أسئلة الاختبار

4. وضوح التعليمات
5. تحسين ظروف التطبيق
6. موضوعية التصحيح

2- الثبات:

يعد مفهوم الثبات من المفاهيم الجوهرية في القياس النفسي، ويمثل مع مفهوم الصدق أهم الأسس، ويتعين توافرهما في المقياس لكي يكون صالحا للاستخدام، والثبات خاصية تتوفر في كل اختبار صادق، بينما الصدق لا يتوفر بالضرورة لكل اختبار ثابت، وذلك بالطريقة نفسها التي نقول بها ان كل مصري افريقي وليس كل افريقي مصري بالضرورة.

تعريف الثبات:

يشير الثبات إلى الدرجة الحقيقية التي تعبر عن أداء الفرد على اختبار ما، ومعنى ثبات الدرجة أن المفحوص سيحصل عليها في كل مرة يختبر فيها سواء بالاختبار نفسه أو بصورة مكافأة له تقيس الخاصية نفسها. سواء اختبر في الظروف نفسها أو في ظروف مختلفة، لا تتدخل فيها عوامل عشوائية غير أن الواقع - الملاحظ - هو أن أداء الفرد يتذبذب ارتفاعا وانخفاضا من وقت لأخر، ومن صورة إلى أخرى من صور الاختبار بل على امتداد الاختبار الواحد، ومن هنا تظهر ضرورة ايجاد السبل للكشف عن حجم هذا الأداء الحقيقي " أي التباين الحقيقي "

وتحديد الدرجة أو النسبة الزائفة من الدرجة " أي تباين الخطأ " الذي يرجع إلى العوامل عشوائية متعددة تؤدي إلى تذبذب الأداء.

وحتى نتمكن من التعرف على الدرجة الحقيقية أو التباين الحقيقي للمقياس، فلا بد من حساب ثباته، ويشير مفهوم الثبات إلى نسبة التباين الحقيقي في الدرجة المستخلصة من اختبار ما إلى التباين الكلي للدرجة على الاختبار.

الثبات إذا هو النسبة من تباين الدرجة على المقياس التي تشير إلى الأداء الفعلي للمفحوص ويتضمن هذا المعنى امكانية تصنيف الدرجة على المقياس إلى مكونين رئيسيين هما: الدرجة الحقيقية والدرجة الخطأ أو الزائفة الناتجة عن عوامل الصدفة العشوائية ولتسهيل المناقشة علينا أن نضع هذه الرموز أمامنا خلال السياق التالي:

ك = الدرجة الكلية على الاختبار التي يحصل عليها المفحوص

ح = الدرجة الحقيقية المعبرة عن أداء الفرد

خ = الدرجة الخطأ أو الجزء الخطأ في الدرجة الناتج عن أخطاء القياس المختلفة.

معنى هذا أن الدرجة الكلية تتكون من جزء حقيقي، وجزء خطأ وهو ما يمكن صياغته في المعادلة التالية:

ك = ح + خ

ويعتمد منطق الثبات على قدرتنا على معالجة العلاقات بين المكونات الثلاثة - التباين والدرجة الحقيقية والدرجة الخطأ، بصورة تمكننا في نهاية الأمر من الوصول إلى ثبات الاختبار.

معامل الارتباط والثبات:

ان الثبات في حقيقة الأمر عبارة عن النسبة من التباين الكلي التي يكون مصدرها الأداء الحقيقي على الاختبار أي المقدار الحقيقي من قدرة المفحوص دون زيادة ناتجة عن ظروف مؤاتية أو نقصان ناتج عن ظروف معوقة.

ولأننا نتوصل إلى تقدير لهذا التباين الحقيقي بحساب الارتباط بين العينة نفسها من هذا الأداء على الاختبار سواء كانت عبر مرحلة زمنية محددة، أو في الوقت نفسه عبر صورتين متكافئتين للاختبار، أو عبر نصفي البنود على الاختبار الواحد، يصبح من الضروري أن نربط بين الثبات ومفهوم الارتباط.

فرغم اننا نحسب ثبات الاختبار بوصفه وحدة إلا إننا نعالجه وفقا لمفهوم الارتباط بوصفه متغيرين هما هاتان العينتان من الأداء المستقل سواء في مرتي اختبار أو تنصيفا أو تجزئة للاختبار.

ومعامل الارتباط في اساسه مقياس للعلاقة بين متغيرين وهو يقدم تقديرا لمدى اتساق التغير في هذين المتغيرين وقوة هذا الاتساق:

1- فمن الناحية الأولى يقدم معامل الارتباط تقديراً لما اذا كان هناك تلازم في التغير أم لا، وقوة هذا التلازم ان وجد.

2- ويقدم من ناحية اخرى تقديراً لإتجاه هذا التلازم في التغير، وما اذا كان إيجابياً بمعنى أن المتغيرين يتزايدان معا ويتناقصان معا، أم أن الزيادة في أحدهما يناظر نقصان في الآخر، وفي أي من هاتين الحالتين لا توجد فروض متعلقة بالعلية في العلاقة بين المتغيرين، فلا يعني تزايدهما أو تناقصهما معاً أن أحدهما سبب والآخر نتيجة، فقد يكون الاثنان متلازمين بهذه الصورة لأنهما معا نتيجة لمتغير ثالث أو متغيرات أخرى تلعب دور السبب.

يعد الثبات من العوامل المهمة أو الخصائص الواجب توافرها لصلاحية استخدام أي اختبار أو قياس، فالمقياس الثابت سوف يعطي نفس النتيجة تقريبا لنفس الشخص عند إجراء القياس لمرات عديدة في نفس اليوم أو أيام مختلفة حيث تكون تلك النتيجة مؤشرا جيدا لقدرات هذا الشخص.

ويتعلق الثبات بدقة القياس بصرف النظر عما يقاس وتتضمن جميع القياسات العملية بعض الخطأ العشوائي الذي يؤدي لعدم ثبات النتائج. والثبات يعني:

- إن الاختبار موثوق به ويعتمد عليه، كما يعني الاستقرار، أي أنه لو أعيد تطبيق الاختبار نفسه على الفرد الواحد فإنه يعطي شيئا من الاستقرار في النتائج.

- وهو اتساق الدرجات التي يحصل عليها نفس الأفراد، في مرات الأجراء المختلفة

- وهو الاختبار الذي لو أعيد تطبيقه على نفس الأفراد فانه يعطي نفس النتائج أو نتائج متقاربة.

ويقاس الثبات إحصائيا بحساب معامل الارتباط بين الدرجات التي حصل عليها الأفراد في المرة الأولى وبين نتائج الاختبار في المرة الثانية. وهو نسبة التباين الحقيقي الداخل في تباين الدرجات التجريبية. ويمكن إن نستدل من صدق الاختبار على انه ثابت، في حين إن الاختبار الثابت ليس بالضرورة إن يكون صادقا.

ويعتبر الثبات من أهم خصائص الاختبارات والمقاييس النفسية بعد الصدق.لأن الصدق أكثر شمولية من الثبات، بناء على ما يقال بأن كل اختبار صادق هو ثابت والعكس غير صحيح بالضرورة. الا أن هذا القول لا يعني الاستغناء عن الثبات، إذا ما تحقق صدق المقياس. وذلك لإننا لا نمتلك أدلة قاطعة على صدق الاختبار. لذا ينبغي تقدير الثبات فضلا عن تقدير الصدق.

والثبات يعني التعرف على العلاقة بين الاستجابة الحقيقية للفرد التي ينبغي الوصول اليها وبين استجابة الفرد على المقياس. كما يعني أيضا الاتساق في النتائج والاستقرار، أي لو كررت عمليات قياس الفرد الواحد لأظهرت درجته شيئا من الاستقرار. كما يعني الموضوعية، بمعنى أن الفرد يحصل على نفس الدرجة كائنا من كان الاخصائي الذي يطبق عليه الاختبار أو يصححه.

ويوفر معامل الثبات المؤشرات الاحصائية للظاهرة موضوع للبحث. والتي من خلالها يحكم على دقة المقياس. فضلا عن ما يزوده للباحث من معلومات

اساسية للحكم على نوعية أسلوب استعمال الاختبار ومدى دقته واتساقه، مع ما يقدمه من نتائج وبيانات عن الظاهرة. ولا تصل هذه الدقة إلى أن يكون معامل الثبات تاما أي يصل إلى درجة مقدارها (1). لأن أي اختبار عقلي أو نفسي أو تحصيلي لا يمكن أن يصل إلى درجة الثبات التامة بسبب عدم امكانية تجنب كل الاخطاء في عمليات القياس. (Ebel،1972)

وهذا يعني أن الدرجة التي تشير إلى ثبات الاختبار، لا تعبر عن الأداء الحقيقي للفرد. بل تمثل الأداء الحقيقي مضافا اليه أخطاء القياس على الدرجة. أي ان درجة الفرد على الاختبار تعبر عن التباين الحقيقي للفرد True Veriance وتباين الخطأ Error Veriance. ولذلك يؤكد جيلفورد Guilford على ضرورة حساب ثبات الاختبار كي تحدد الدرجة الحقيقية أو التباين الحقيقي للاختبار.اذ ان معامل الثبات يوضح نسبة التباين الحقيقي في الدرجة المحسوبة على الاختبار. (Guilford،1954،P.260) (فرج،1980 ص333).

وتشير انستازيAnastiasi إلى أن مصدر الثبات هو الاتساق في الدرجات، التي يتم الحصول عليها من نفس الأفراد اذا ما اعيد عليهم الاختبار نفسه في اوقات مختلفة، أو مع مجاميع مختلفة. وهذا المفهوم للثبات يركز على حساب خطأ القياس لدرجة الفرد.(Anastiasi،1976،P.71) ويتفق سكس sax مع انستازي في التأكيد على خطأ القياس، حيث يرى أن درجة الفرد على المقياس إنما هي الدرجة الحقيقية له.مضافا اليها درجة الخطأ. حيث تعني الدرجة الحقيقية متوسط افتراضي لعدد غير محدد من تطبيق الاختبار عند الغاء التاثيرات

المصاحبة عند التطبيق.تلك التاثيرات الناتجة عن الخبرات المتراكمة لدى من يعاد تطبيق الاختبار عليهم. (Sax، 1989،p.260)

ويشير ساكس Sax أيضا إلى أن خطأ القياس أو درجة الخطأ هي نتيجة الاختلاف بين الدرجة التي يحصل عليها الفرد على المقياس (الدرجة المحققة أو الدرجة الملاحظة) والدرجة الحقيقية. وعندما يكون خطأ القياس صفرا فإن الدرجة التي يحصل عليها الفرد على المقياس تكون هي الدرجة الحقيقية. وعليه فإن الدرجة الملاحظة تساوي الدرجة الحقيقية (الصحيحة) + درجة الخطأ: Obtained Score =true Score + error Score

ولما كان من الصعوبة بمكان الوصول إلى الدرجة الحقيقية بصورة مباشرة فان بالإمكان الوصول إليها عن طريق درجة الخطأ. حيث أن الانحراف المعياري لدرجة الخطأ يمكن الوصول اليه عن طريق الخطأ المعياري للقياس. (Sax،1989،pp.261-262) ويشار إلى الدرجة الحقيقية بالتباين الحقيقي ودرجة الخطأ بتباين الخطأ.

وعلى العموم يمكن الوصول إلى ثبات الاختبار على انه الارتباط بين الاختبار ونفسه من خلال المعادلة الاتية:

معامل الثبات: التباين الحقيقي =التباين الكلي- تباين الخطأ. (فرج،1980)

وتبقى أهمية الثبات بارزة في كل اختبار. ومن المؤشرات التي توضح تلك الاهمية:

1. إن الثبات ضروري. ولكنه ليس الحالة التي تحدد دقة الاختبار.ونوعيته حيث قد تكون تلك الدرجات العالية تقيس شيئا آخر لا علاقة له بالظاهرة المراد قياسها. أما إذا كانت درجة الثبات منخفضة فهذا يعني أن الاختبار قد فشل في قياس الظاهرة أو أن هناك خللا ما.

2. إن الاختبار أو المقياس هو مجموعة من الفقرات. لا ثبات له بحد ذاته. وإن الثبات لا يظهر إلا حين يطبق الاختبار على عينة الأفراد.

3. إن الفروق بين درجات الفرد، اذا ما كانت قليلة قياسا إلى الفروق بين درجات الأفراد الآخرين، فان الاختبار يتجه إلى اعطاء ثبات عال. أما إذا كانت تلك الفروق بين درجات الفرد ذات صلة كبيرة بالفروق بين الأفراد فان الثبات سيكون منخفضاً. (Ebel, 1979).

ويشير ثورندايك إلى ان هناك عدة عوامل يمكن ان تؤثر على معامل الثبات، منها:

1..سمات الفرد وقدراته ومهارته، ومنها مستوى قدرة الفرد ومهارته وطريقته في حل الاختبار وقدرته على فهم التعليمات.

2. مستوى قدرة الفرد فيما هو مطلوب منه لأداء الاختبار والمعلومات والمهارات الخاصة بنوع معين من مفردات الاختبار. فضلا عن انواع معينة من مفردات الاختبار وأثر عوامل الصدفة والتخمين لدى بعض الأفراد لنوع معين من المفردات.

3. العوامل التي تؤثر في الظاهرة، كالتعب والدافعية والتـوتر الانفعـالي. ومنطـق الاختبـار نفسه.وفهم طريقة اداء الاختبار والظروف الفيزيائية المؤثرة على الاختبار كالهدوء، ودرجة الحرارة ونوع الاضاءة وراحة مكان الجلوس، وما إلى ذلك.

4. العوامل النوعية الخاصة بالاختبار مثل فهـم الأعـمال المطلوبـة مـن الاختبـار والتأهـب العقلي الوقتي لحل اختبار ما.وكذلك العوامل الخاصة ببعض بنـود الاختبـار وعـدم الدقـة والتركيز، وعوامل الصدفة، ووضوح التعليمات، والعوامل المؤثرة في الأداء، والتبـاين الممثـل بعامل الخطأ.(الامام وآخرون، 1999)

وعلى العموم فان الحد الادنى المقبول لمعامل الثبات يتوقف على الهدف من القياس وعلى مسـتوى الدقـة المـراد تحقيقـه. ولا توجـد معـايير مطلقـة متفـق عليهـا لتقـويم ثبـات الاختبارات. إلا أن كيلي Kelley قد توصل بناء على افتراضا ت احصـائية معينـة إلى حـدود دنيا لتقويم معاملات الثبات لبعض أهداف القياس. فاذا كان الهدف تقويم مستوى الفرد، فيجب أن لا يقل معامل الثبات عن (0.94) اما اذا كان الهدف دراسة الفروق بين الصفات المختلفة (بروفيل) فيجب أن لا يقل الثبات عـن(0.90) للمجموعـات و(0.98) للأفـراد. إلا أن هذه القيم تعد مستويات مثالية قلما تتحقق عمليا وخاصـة في اختبـارات الشخصـية (أبو علام، 1987)

وهذا القول فيه مغالاة كبيرة. حيث أن المقياس أو الاختبار يختلف باختلاف طبيعة الظاهرة ومدى ثباتها النسبي. ومن ثم فلا يمكن تحديد حدود دنيا لعدم تحديد ثبات الظاهرة كما هي على الطبيعة قياسا إلى غيرها من الظواهر. فظاهرة القلق مثلا هي غير ظاهرة الثقة بالنفس، أو الانجاز أو الدافعية. ومن ثم لاحدود دنيا مقبولة للثبات. وأن ما يلجأ إليه بعض الباحثين من اختبار دلالة معامل الثبات لا معنى له. لأن معامل الثبات هو ليس معامل الارتباط. مع أن معامل الارتباط يستعمل للوصول إلى معامل الثبات. وأن الطريقة الصحيحة لاختبار معامل الثبات هو موازنته بالدراسات السابقة التي تناولت نفس الظاهرة. أو أن نربع معامل الثبات فاذا ظهر انه اكثر من (0.50) فيمكن أن يكون ذلك مؤشرا على الثبات يمكن الاعتماد إليه في غياب الدراسات السابقة.

اما فيما يتعلق بالطرائق المتبعة للوصول إلى معامل ثبات الاختبار فهي كما هو معروف، طريقة اعادة الاختبار وطريقة التجزئة النصفية وملحقاتها وطريقة الصور المتكافئة، وطريقة تحليل التباين وانواعه بما فيها معامل الموضوعية بالنسبة للمصححين والممتحنين. ولكل طريقة لها ما لها وعليها ما عليها. ويمكن تلخيص انواع معاملات الثبات وطرق تقديرها وتباين الخطأ المتوقع في الآتي:

- اولاً: معاملات الثبات:

1. معامل الاستقرار.

2.معامل التكافؤ.

3. معامل الاستقرار والتكافؤ معا.

4. معامل الاتساق الداخلي.

5.معامل الاتساق داخل الاسئلة.

6. معامل الموضوعية.

- ثانياً: تباين الخطأ:

1.التغيرات الموقتة.

2.عينة محتوى الاختبار.

3.التغيرات المؤقتة وعينة محتوى الاختبار.

4.عدم تجانس نصفي الاختبار.

5.التجانس الكلي لاسئلة الاختبار.

6.ذاتية المصحح وذاتية الممتحن والتغيرات الموقتة.

ان ثبات الاختبار يجب أن يكون على درجة عالية من الدقة والاتقان والاتساق، فيما يزودنا به من بيانات، إذ إن مفهوم الثبات من المفاهيم الأساسية التي يجب أن تتوفر في الاختبار لكي يكون صالحا للاستخدام، ففي كل اختبار يوجد قدر من أخطاء القياس، وقد تكون الأخطاء قليلة أو كثيرة، مما تؤثر في نتائج القياس،

وتسمى هذه الاخطاء بأخطاء الصدفة، إذ إنه لا يوجد اختبار سواء كان تحصيليا أو عقليا أو نفسيا يحصل على درجة ثبات كاملة، لأن من غير الممكن التخلص من الاخطاء في الاختبار، لإننا عندما نريد استخراج معامل الثبات، نحسبه بناءً على ارتباط درجات المرة الثانية بنفس الصورة أو بصورة متكافئة من الاختبار، وهنا لابد أن تكون معرض للخطأ، وهذا الخطأ هو أن هؤلاء الأفراد انفسهم يتغيرون، وعلينا أن نعرف مدى تغيرهم حتى نعرف سبب تغير درجاتهم في المرات المختلفة، فقد لا يكون التغير في الاختبار بل فيمن يطبق عليهم الاختبار، اذ ان الدرجة التي يحصل عليها الفرد في أي اختبار لاتعبر عن الأداء الحقيقي فقط وانما تمثل الأداء الحقيقي للفرد مع عوامل الخطأ، أي بمعنى ان درجة الفرد على الاختبار تعبر عن التباين الحقيقي للفرد وتباين الخطأ، وبذلك فان درجة الاختبار تتضمن قدرا من الخطأ، قد يؤدي إلى زيادة الدرجة الحقيقية أو إلى نقصان في الدرجة عن ما يستحقه الفرد المقاس، مما يؤدي إلى تقدير أداء الفرد أقل أو اكثر من الواقع الذي يستحقه، وأفضل طريقة لمقارنة هذه الدرجات هي حساب معامل ارتباط درجات الاختبار في المرة الاولى بدرجات نفس الاختبار في المرة الثانية بنفس الأفراد، وعندما تكون درجات المرتين نفسها متسقة فان ذلك يدل على ثبات اجاباتهم ويستخدم في حساب معامل الثبات. ومعامل الارتباط بين درجات الأفراد في المرتين ويطلق على النتيجة التي نحصل عليها بمصطلح (معامل الثبات) ويتراوح معامل الثبات بين درجتين (صفر و 1) ويعد الصفر ادنى معامل ثبات، اما الدرجة (1) فتمثل أعلى معامل ثبات، ومن الصعب الوصول إلى معامل ثبات يساوي (1)، وذلك بسبب الاخطاء العديدة

التي تتعلق بنتائج الاختبار والتي لا تخضع للضبط العلمي والتحكم الدقيق، وكذلك الحالة النفسية أو الجسمية للفرد والحالة الفيزيقية وغيرها مما تؤثر مباشرة في نتائج الثبات.

إن ثبات الاختبار يعني ان يعطي الاختبار نفس النتائج إذا ما أعيد على نفس المجموعة في نفس الظروف وبمعنى آخر لو كررت عمليات قياس الفرد الواحد، لأظهرت الدرجة شيئا من الاتساق أي أن درجته لاتتغير جوهريا بتكرار إجراء الاختبار.

أي أن مفهوم الثبات يعني أن يكون الاختبار قادرا على أن يحقق دائما النتائج نفسها في حالة تطبيقه مرتين على نفس المجموعة.

طرق الحصول على ثبات الاختبار:

تتعدد أساليب حساب الثبات ويختص كل أسلوب منها بتقدير نوعية محددة من " تباين الخطأ " وهو التباين العشوائي والدخيل الذي يؤثر في ثبات القياس الذي نحصل عليه في كل مرة نستخدم فيها مقاييسنا المختلفة، وتصلح بعض الأساليب لحساب ثبات مقاييس معينة بينما لا تصلح لحساب ثبات مقاييس أخرى ولا يعني هذا أن المقياس الواحد لا تصلح إلا طريقة واحدة لحساب الثبات ففي بعض الحالات يمكن أن نستخدم أكثر من أسلوب ثبات لاختبار الواحد بهدف التعرف على مصادر " تباين الخطأ " الذي يؤثر في استقرار أو اتساق الدرجة التي نحصل عليها منه.

يتعين دائما دراسة أي الأساليب أصلح لحساب ثبات الاختبار والمحك، لكون أسلوب ما أفضل من أسلوب اخر هو أن الأسلوب المناسب يؤدي دائما إلى الحصول على أعلى معامل ثبات ممكن.

والنقطة المهمة التي يجب أن تكون واضحة في الاذهان، هي أن الاختبار نفسه بوصفه مجموعة من البنود أو الاسئلة أو الأعمال المعينة لا ثبات لها. وان الثبات هو ثبات الإجابة عن الاختبار وثبات الأداء على الاختبار.

طرق حساب معامل ثبات الاختبار:

اولاً: طريقة الصور المتكافئة:

تعد هذه الطريقة من أفضل الطرق ملائمة مع الاختبارات التحصيلية ويفترض تكوين صورتين متكافئتين من الاختبار الواحد بحيث تكون هذه الصور متكافئة، وتتوفر فيها نفس مواصفات الاختبار الذي نريد التحقق من ثباته بحيث يحتوي على نفس العدد من الأسئلة وأن تكون صياغة الاسئلة متماثلة ودرجة الصعوبة واحدة، وأن تتضمن محتوى واحد، وأن تتفق معه في جميع المظاهر الأخرى مثل التعليمات والأمثلة والشكل العام.

فمثلا وجود صورتين من الاختبارين للقراءة يجب أن يتضمنا نصوصا وأسئلة لها نفس الصعوبة، ويسأل فيها نفس النوع من الأسئلة، أي ان يكون هناك توازن واحد بين الاسئلة، وكذلك يجب أن تمثل فيها نفس نوع النصوص سواء كانت وصفية أو حوارية، ولكن النصوص الخاصة ومضامين الاسئلة يجب أن تكون

مختلفة، فإذا أصبحت لدينا صورتان من الاختبار، فيمكن أن نطبق الصورتين، فإما أن يعطي الواحد بعد الآخر مباشرة في نفس الوقت إذا لم يكن هنالك اهتمام بالاستقرار عبر الزمن، وبعد ذلك يحسب معامل ارتباط بين درجات الاختبارين فيكون بذلك معامل ثبات تكافؤ، أو أن يطبق الاختبارين بعد فترة زمنية، أي أن تكون هنالك فترة مناسبة بين أجراء صورتين، وعند ذلك يكون معامل الثبات المحسوب بهذه الطريقة هو معامل التكافؤ والاستقرار. ان طريقة الصور المتكافئة تقدم أساسا سليما جيداً لتقدير الدقة في الاختبار إلا أن هذه الطريقة تثير عدد من المشكلات العملية إذ أنها تتطلب توفر صورتين متكافئتين، ففي بعض الاختبارات لايمكن إعداد صورة للاختبار أو قد لاتوفر الوقت لإجراء الاختبار الثاني. كما أن عامل أثر التدريب والألفة بالاختبار يزداد كلما اقتربت الصورة من الأصل، مما يؤثر على ثبات الاختبار.

إن استخدام صورة متكافئة متعاقبة يعني تقديم الاختبار نفسه بعد فترة ولكن ببنود مختلفة، حيث يكون الاختلاف دائما في صياغة البنود. وعلى ذلك يمكن أن نجد طريقتين من طرق استخدام الصور المتكافئة:

1- الصور المتكافئة الفورية.

2- الصور المتكافئة المتعاقبة.

ثانياً- طريقة إعادة الاختبار:

تتطلب هذه الطريقة اعادة تطبيق الاختبار مرة أخرى على أفراد المجموعة نفسها بعد فترة زمنية ملائمة، ثم تحسب بعد ذلك معامل الارتباط بين الدرجات التي حصل عليها أفراد العينة في المرة الاولى والثانية، ويسمى معامل الارتباط المستخرج بهذه الطريقة بمعامل الاستقرار، أي استقرار نتائج الاختبار خلال الفترة بين التطبيقين للاختبار

ومن الضروري عند حساب الثبات بطريقة إعادة الاختبار تقدير الوقت بين التطبيقين، فإن هنالك احتمال كبير من تدخل عامل النمو الجسمي والعقلي لدى الفرد المقاس، وعموما فليس هناك اتفاق على طول الفترة الزمنية بين تطبيقي الاختبار، ولكن يمكن القول بصورة عامة، إن الفترة القصيرة تعتمد على نوع الظاهرة المقاسة أي عدم تأثر الاستجابات في التطبيق الثاني بتذكر الاستجابات في التطبيق الأول، وأن أفراد العينة من الأطفال حتى لا يدخل عامل النمو الجسمي والعقلي. أما إذا كانت السمة المقاسة لا تتأثر بالنمو وأفراد العينة من الكبار ففي مثل هذه الحالة فان الفترة الطويلة تكون ملائمة لها، إن طريقة إعادة الاختبار لاتصلح لحساب ثبات الاختبارات التحصيلية، وتكون أكثر ملائمة مع الاختبارات الشخصية.

ومن المآخذ على هذه الطريقة ان الدرجات التي يحصل عليها الأفراد في المرة الثانية من تطبيق الاختبار تكون أعلى بقليل من درجاتهم في التطبيق الأول، وذلك بسبب الألفة، وتذكر الإجابة كما أن الموقف التجريبي قد يختلف في المرة

الأولى عن المرة الثانية، مما تتأثر النتائج النهائية باخطاء القياس التي يصعب إخضاعها للظروف التجريبية، كما أن هذه الطريقة تكلف جهدا ووقتا. كما أن هناك عوامل كثيرة تتدخل في تحديد الفترة بين الاختبار وإعادة الاختبار، وبصفة عامة يتعين ألا تقل هذه الفترة عن أسبوع ولا تزيد عن ستة أشهر، ويحدد مضمون الاختبار هذه الفترة الزمنية كما تحددها طبيعة العينات المستخدمة في الدراسة.

- كما يتأثر اختيار الفترة بين الاختبار واعادته بمضمون الاختبار.

- ويتأثر أيضا من حيث طبيعة العينة المستخدمة في دراسة الثبات.

- وإذا كانت مادة الاختبار من النوع الذي يسهل تعلمه أو تذكره أو الاستفادة فيه من التدريب أو يتعرض الأفراد فيه لنوع من الألفة فمن الأفضل ان تطول الفترة بين الاختبار وإعادة الاختبار.

- وإذا كانت العينات من الاطفال صغار السن أو الاشخاص الذين يعبرون مراحل ارتقائية سريعة وحاسمة أو يعالجون بعقاقير تؤثر في ذاكرتهم أو انتباههم أو وظائفهم المختلفة فمن الأفضل أن تنخفض الفترة بين الاختبار وإعادة الاختبار.

- إن انتخاب الأسلوب المناسب، والفترة المناسبة لإعادة الاختبار، وحدود التجانس المطلوب في العينة يعتمد في الجانب الأكبر منه على تحليلنا

السيكولوجي لما يقيسه الاختبار وكيفية قياسه، كما يعتمد على ما يوفره لنا التراث عن هذه الوظائف، ونموها أو ارتقائها وتأثرها بالمتغيرات المختلفة.

ثالثاً- طريقة التجزئة النصفية:

لاتحتاج هذه الطريقة إلى تطبيق الاختبار مرتين، ولا إلى تصميم صورتين متكافئتين للاختبار، وإنما تتطلب تصميم اختبار متكافئين، على أن يحتوي القسم الأول على الدرجات الفردية 1، 3، 5، 7، والقسم الثاني على الدرجات الزوجية 2، 4، 6، 8.وتمتاز هذه الطريقة بأنها تتغلب على العوامل المؤثرة في أداء الأفراد مثل الوقت والجهد والتعب، كما أنها تمتاز بتوحد ظروف الإجراء توحيدا كاملا. ومن الشروط الواجب اتباعها في تصميم الاختبار عند استخدام هذه الطريقة ان تتكافأ البنود الفردية مع البنود الزوجية. وللحصول على الثبات نستخدم معاملات الارتباط بين درجات كل فرد على البنود الفردية والزوجية، ويستخدم بعد ذلك معادلة سبيرمان براون.

- غير ان هناك بعض المقاييس لا تقبل القسمة النصفية المتساوية لاعتبارات كثيرة:

* فاذا كان الاختبار ينقسم إلى اختبارات فرعية يقيس كل منها وظيفة معينة، فإن قسمته في هذه الحالة خاطئة.

* يضاف إلى ذلك أن بعض الاختبارات تقيس مجرد سرعة الأداء، وهنا لا يمكننا أن نفترض أن النصف الاخير من الاختبار يتوفر للمفحوص فيه وقت مساو لما يتوفر له في النصف الأول من الاختبار.

- الأساس في الاختبارات التي تقبل استخدام أساليب التنصيف هو أن يكون تصميمها قائما على توفير وحدات متجانسة بقدر الامكان، ولا يكون الأداء عليها تراكمي، بمعنى أن كل خطوة تترتب على سابقتها، وتبنى كل اجابة على الإجابات الواقعة قبلها في الترتيب، وألا يكون الأداء مطلوبا خلال مدى زمني يتطلب السرعة في العمل، والتوقف عنه تحت محك انقضاء الوقت.

وهنالك عدة طرق لتقسيم الاختبار منها:

أ- القسمة النصفية: وذلك بقسمة الاختبار إلى نصفين متساويين فمثلا لو كان لدينا اختبار يتكون من (60) فقرة، فسيكون النصف الاول (1-30) فقرة والثاني من (31-60) فقرة.

ب- الفردي والزوجي: يتم ذلك بقسمة فقرات الاختبار إلى نصفين بحسب أرقام الفقرات، فالفقرات ذات الارقام الفردية مثلا تعد قسما والفقرات ذات الأرقام الزوجية قسما آخر.

ج- التقسيم الجزيئي: تشمل بعض الاختبارات على اختبارات فرعية، وفي مثل هذه الحالة لايمكن استخدام أي نوع من الأنواع اعلاه، لذلك يقسم الاختبار إلى جزئين أو ثلاثة اجزاء، بحسب ما يحتويه الاختبار من جوانب فرعية، فمثلا إذا

كان الاختبار من (6) اختبارات فرعية، فيكون القسم الأول من الفروع (1-3-5) والقسم الثاني من الفروع (2-4-6) وهكذا.

إن معامل الارتباط المستخرج بين نصفي الاختبار، يفسر بأنه معامل الاتساق الداخلي، ولما كان معامل الثبات لايقيس التجانس الكلي للاختبار، لأنه ثبات لنصف الاختبار، فهنالك طرق متعددة لتلافي هذا النقص، وذلك باستخدام معامل سبيرمان براون ومعادلة رولون وطريقة كودر – ريتشردسون وغيرها.

متى نقوم بالتنصيف:

عندما نختار أحد أساليب التنصيف التي أشرنا اليها لحساب ثبات اختبار معين فإننا لا نقوم بتنصيف فعلي للاختبار قبل تطبيقه، بل يقدم الاختبار بكل بنوده أو بكل أجزائه إلى عينة المبحوثين المستخدمة لتقدير ثبات الاختبار، وبعد أن يتم اختبارهم نقوم بإجراء هذه القسمة عن طريق تسجيل الدرجة على النصف أو الجزء الأول على حده لكل فرد، ثم تسجل الدرجة على النصف أو الجزء الثاني على حدة للأفراد أنفسهم أو الدرجة على البنود الفردية على حدة، وفي مقابلها الدرجة على البنود الزوجية.

معنى هذا أننا لا نقوم بتنصيف الاختبار نفسه، بل بتنصيف أداء المفحوصين على النصفين وحساب الارتباط بين أدائهم على النصفين، كائنة ما تكون طريقة التنصيف.

العوامل المؤثرة في الثبات:

جميع طرق حساب الثبات كمية إحصائية، ولا يمكن تحديد مستوى الثبات قبل تطبيق الاختبار على عينة تجريبية. لكن هناك عوامل يمكن أن تساهم في رفع الثبات ينبغي أن يأخذها معد الاختبار ومستخدمه في الاعتبار:

1- طول الاختبار: إن عدد أسئلة الاختبار عامل مؤثر في درجة ثبات الاختبار، فكلما زاد عدد الفقرات أدى ذلك إلى ارتفاع معامل الثبات.

2- زمن الاختبار: يتأثر ثبات الاختبارات الموقوتة بالزمن المحدد لها، وقد أكدت أبحاث ليند كريست، وبذلك يزداد الثبات تبعا لزيادة الزمن حتى يصل إلى الحد المناسب للاختبار فيصل إلى نهايته العظمى، ثم يقل بعد ذلك كلما زاد الزمن عن ذلك الحد.

3- غموض الأسئلة: عندما تكون الأسئلة غامضة وطويلة تقلل من الثبات، والاسئلة الواضحة والموضوعية والقصيرة تزيد من الثبات، وهذا يتطلب من الباحث الدقة في اختيار الألفاظ والعبارات ونوعها، عند صياغة الاسئلة حتى يجعلها قادرة على الوصول إلى الثبات الحقيقي.

4-صعوبة الفقرات ودقتها: ان وجود فقرات صعبة في الاختبار لا يستطيع أغلب الأفراد أو جميعهم من الاجابة عنها تؤثر في الاختبار من حيث ثباتها، إلا أن حذفها لا يؤثر على درجة المفحوص، كما أن وجود فقرات سهلة يستطيع جميع أفراد العينة الإجابة عنها تؤثر في معامل ثبات الاختبار، إلا أن حذفها لا

تـؤثر عـلى الاختبـار لـذلك ينبغـي حـذف الأسـئلة الصـعبة أو السـهلة لرفـع ثبـات الاختبار.

5- حالة الفرد: يتأثر الثبات بحالة الفرد النفسية والصحية ومـدى تدربـه عـلى الموقـف الاختبـاري، فـالمرض والتعـب والتـوتر الانفعـالي قـد يـؤدي إلى نقصـان الثبات.

6- ظروف إجراء الاختبـار: إن أي تغيـير في الظـروف الخاصـة بـإجراء القيـاس مـن الاختبار لآخر يؤثر على نتائج المقياس، ويحد من عوامل الخطأ الذي يؤثر في ثبـات الاختبار.

7- موضوعية الاختبار: تعد موضوعية التصحيح من بين العوامل المـؤثرة في ثبـات لاختبـار، وعلى الاخص في الاختبارات التي تعتمد على تقدير المصحح كاختبـارات المقـال التحصـيلية واختبارات الإبداع، حيث أن تباين التصحيح يؤثر في زيادة تباين الخطأ، وبالتالي إلى نقصان معامل الثبات.

8- التخمين: يلجأ بعض من المفحوصين في حالة عـدم تأكـدهم مـن الاجابـة الصـحيحة إلى التخمين، مما يؤدي إلى خفض ثبـات الاختبـار، حيـث زيـادة اثـر التخمـين يقـود إلى نقـص الثبات.

9- تباين العينة وتجانسها: يرتبط الثبـات بتبـاين الاختبـار فيـنقص ثبـات الاختبـار عنـدما يــنقص التبــاين ويــزداد تبعــا لزيــادة التبــاين، فالســهولة والصــعوبة تؤديــان إلى

التجانس وبالتالي إلى خفض الثبات، والاسئلة المتدرجة في الصعوبة تؤدي إلى رفع الثبات.

ويمكن أن نجعل الاختبار ثابتا عن طريق:

أ- جعل الاختبار طويلا بدرجة معقولة، حتى يؤدي إلى التخلص من عامل الصدفة أو التقليل منه.

ب- العمل على تصحيح الاختبار بطريقة موضوعية ما أمكن.

ج- كتابة تعليمات الاختبار بدقة ووضوح، حتى لا يقع الطالب في أخطاء تعود بالدرجة الأولي إلى غموض أو عدم فهم التعليمات.

10- حجم عينة الثبات:

يجب ألا يقل حجم عينة الثبات عن 30 مفحوصا، ويحكم هذا الحد الأدنى اعتبارات احصائية مهمة، اذ تعد العينات التي تقل عن ثلاثين مفحوصا عينات صغيرة تستخدم في معالجاتها الاحصائية أساليب تختلف عن اساليب معالجة العينات الكبيرة التي يفترض فيها اعتدالية التوزيع.

11- الخطأ المعياري للقياس:

مقياس للخطأ المتوقع في درجة الاختبار، أنه المقدار الذي في حدوده تزيد أو تنقص الدرجة المشاهدة عن الدرجة الحقيقية للطالب. وكلما قل خطأ القياس المعياري كلما زادت الدقة في درجات الاختبار، وخطأ القياس المعياري هو عبارة

عن الانحراف المعياري لمجموعة من الفروقات كل منها يساوي الفرق بين الدرجة الحقيقية والدرجة الحاصل عليها. الخطأ المعياري له علاقة مباشرة بثبات الاختبار، فكلما كانت درجة ثبات الاختبارعالية كلما قل خطأ القياس المعياري، ويمكن تعريف الخطأ المعياري للقياس أيضا بأنه الانحراف المعياري لدرجات طالب ناتجة عن تكرار إجابته للاختبار نفسه (أو اختبار مماثل له)، ولأن مثل هذه البيانات لا تتوافر عادة فإنها تؤخذ من اختبار مجموعة واحدة من الطلاب.يمكن حساب الخطأ المعياري للاختبار بمعرفة تباين (variance) الدرجات، وثبات الاختبار.فإذا حصل طالب على درجة مقدارها 90 في اختبار خطأه المعياري يساوي 3 فإننا نقول أن هناك احتمالا قدره 68% أن تكون الدرجة الحقيقة للطالب واقعة في فترة الثقة بين 87 و 93. وغالباً ما نستخدم فترة ثقة عالية جداً، فأكثر فترات الثقة استخداماً هي 95% و 99% وبالتالي تكون نسبة الشك5 0, ، و1 0,. علي التوالي

والمعادلة التالية تبين حساب الخطأ المعياري للقياس:

الخطأ المعياري للقياس = الانحراف المعياري لتوزيع الاختبار × الجذر التربيعي لثبات الاختبار – 1

من خلال ما تقدم يمكن ان يستنتج ما ياتي:

1. أن درجة ثبات المقياس التـي نحصـل عليهـا بـالطرائق المختلفـة هـي ليسـت الدرجـة الحقيقية للثبات. بل هي الدرجة الحقيقية مضافا اليها درجة الخطـأ النـاتج مـن عمليـات القياس المختلفة.

2. أن ثبات أي اختبار هـو ثبـات للتطبيـق عـلى مجموعـة مـا. ولـذا فـإن مـن الواجب أن يستخرج الثبات عند كل تطبيق لكي نزيد مـن قيمـة مـؤشر الثبـات للاختبار.

3. أن الطرق المختلفة لاستخراج معامـل الثبـات هـي مـؤشرات للثبـات. ولـذا يستحسـن الحصول على أكثر من درجة لمعامل الثبات. اذ كلما زادت المؤشرات كلما كان ذلك أفضـل، على أن تتنوع تلك المؤشرات بين استقرار الاستجابة وتجانسها.

4. أن كل أسلوب من أساليب استخراج معامل الثبـات يخـتص بنـوع مـن تبـاين الخطـأ. ونوعية الثبات المراد الوصول إليه، وأن الباحث يمكنه أن يبرر استعماله لأسلوب ما أو اكـثر من أسلوب، وبما يخدم دقة الأداة المستخدمة.

5. أن اختبار دلالة معامل الثبات لا معنى له. وأن معرفة دلالة معامـل الثبـات تـتم مـن خلال موازنته بالدراسات السابقة، وليس باختبار معنوية معامل الارتباط.

العلاقة بين الثبات والصدق:

إن الصدق يتضمن الثبات وهو مظهر من مظاهره، كما أن الصدق أعم وأشمل من الثبات، إذ أن ثبات ودقة الاختبار لاتدل على صدقه، بل من الممكن أن يكون الاختبار ثابتا ودقيقا، وأن درجة الفرد عليه لاتتغير كثيرا من إجراء إلى آخر، ولكنه ليس صادقا بمعنى أن يقيس دقة وثبات عاملا آخر غير العامل الذي صمم الاختبار لقياسه، ومن جهة أخرى فان كل اختبار صادق هو ثابت بالضرورة، ولكن لايمكن القول أن كل اختبار ثابت هو صادق بالضرورة إذ أن الاختبار الصادق يقيس فعلا ما أعد لقياسه، فان درجته معبرة عن الأداء الحقيقي أو القدرة الفعلية للفرد وتعبر عن الوظيفة المقاسة بكل دقة وبالتالي ستكون ثابتة في الوقت نفسه.

أي أن الاختبار الثابت ليس بالضروري أن يكون صادقا ذلك أنه قد يقيس وظيفة أخرى عدا الوظيفة المخصص لقياسها أو يقيس وظيفة أخرى إلى جانب الوظيفة التي خصص لها.

لتحديد معنى الصدق لابد من التمييز بينه وبين الثبات، فالثبات شرط أساسي لتحقيق الصدق، فلا يمكن أن يكون الاختبار صالحا للاستخدام، ولا يمكن أن يقيس الصفة التي يزعم قياسها حتى تتوفر فيه صفة الثبات. والاختبار الذي يحصل فيه الفرد على درجة 70 ونحن في شك من أن درجته الحقيقية قد تكون 75 أو 60 لا يمكن أن تسمى الدرجات فيه صادقة.

الاختبار الذي يضعف ثباته لا يمكن الوثوق به ولا تستخراج دلالات صحيحة من خلاله، ولا تبنى قرارات صائبه بناء عليه.

من جهة أخرى فإن توفر الثبات لا يكفي لتحقيق الصدق، إن وجود ميزان ثابت دقيق لا يبرر لنا استخدامه لقياس الطول أو الحرارة، إذ لا بد من وجود علاقة منطقية أيضا بين محتوى الأداة والصفة المرغوب قياسها.

إذن لا يمكن أن نحكم على صدق المقياس من مجرد ثباته، فالثبات شرط ضروري لكنه لا يكفي وحده ولا يكون الاختبار صادقا حتى يكون ثابتا، لكنه يمكن أن يكون ثابتا ولا يكون صادقا.

3- الموضوعية:

من أهم صفات الاختبار الجيد أن يكون موضوعيا لقياس الظاهرة التي أعد أصلاً لقياسها، والموضوعية هي التحرر من التحيز أو التعصب، وعدم إدخال العوامل الشخصية للمختبر كآرائه وميوله الشخصية وحتى تحيزه أو تعصبه، فالموضوعية تعنى وصف قدرات الفرد كما هي موجودة فعلا، لا كما نريدها أن تكون. وهي عدم اختلاف المقدرين في الحكم على شيء ما أو على موضوع معين، أي أن هناك فهماً كاملاً من جميع المختبرين بما سيؤدونه، وأن يكون هناك تفسير واحد للجميع، وأن لايكون هناك فرصة لفهم معنى آخر غير المقصود منه. ويلاحظ أن جميع المقاييس الموضوعية من ميزان طبي يكون فيها جزا ولو

بسيط ذاتيا، حيث أن الذي قام بتصنيع هذه الأدوات أفراد، ولكن نسبة الذاتية يمكن أن تتلاشى، وإذا حدث أخطاء في القياس تكون غالبا من مستخدم الأداة.

وعليه يجب على كل من يقوم بتطبيق اختبارات بدنية أو مهارية أن يحدد التعليمات لكل اختبار، وأن تكون التعليمات واضحة، ثم القيام بعمل نموذج أمام المختبرين بالإضافة إلى الاطمئنان على صحة الأداة، والأجهزة المستخدمة وأن يثبت جميع الشروط الواجب اتخاذها أثناء التطبيق، بالإضافة إلى تدريب بعض الأفراد من ذوي الخبرة لكيفية استخدام الأدوات والأجهزة، وكيفية استخراج النتائج.

إن موضوعية إجراءات تطبيق أي اختبار يحكم عليها بواسطة درجة الاتفاق بين الدرجة النهائية التي يقدمها ملاحظان مستقلان أو اكثر، وكلما كانت الملاحظة والتقويم ذاتيين، انخفضت درجة الاتفاق بين الحكمين. وفي الاختبارات التي يختار فيها المختبر البديل الصحيح أو البديل الأفضل من بين عدة بدائل، تكون الموضوعية عالية، لان بإمكان المصححين كلهم استخدام مفتاح التصحيح والاتفاق على النتائج كاملا. وعلى العكس من ذلك فإن اختبارات المقال تفسح المجال أمام الاختلاف الواسع بين المصححين.

العوامل التي تؤثر في معامل الموضوعية:

- درجة وضوح الاختبار، فكلما كان الاختبار واضحا للمختبر والمحكمين ارتفعت الموضوعية.

- مدى فهم المختبرين لطبيعة الاختبار، وطريقة تنفيذه، والتسجيل.

أما شروط تحقيق الموضوعية:

1- يجب إيضاح شروط الإجراء والتعليمات بدقة.

2- يجب اختيار المحكمين المدربين على طرق القياس الصحيحة والدقيقة، للحد من التحيز في التقدير.

3- يجب تبسيط إجراءات القياس لضمان الحصول على نتائج دقيقة.

4- استخدام أجهزة قياس حديثة وإلكترونية للوصول إلى أدق النتائج في زمن بسيط.

5- متابعة تنفيذ الاختبار للأفراد المختبرين للتأكد من تنفيذ نفس الشروط والتعليمات، والتسجيل للنتائج.

6- إعداد مفاتيح التصحيح الخاصة بكل اختبار مقدما قبل تطبيقه.

7- اتباع تعليمات الدليل المرفق بالاختبار بدقة لتحديد طريقة التقدير، وذلك للحد من ذاتية الفاحص.

الفصل الرابع

مقدمة:

يعد مفهوم معايير الاختبارات من المفاهيم الأساسية المتعلقة بتفسير دراجات الاختبارات مرجعية الجماعة أو المعيار، فالدرجة التي يحصل عليها فرد في اختبار ما والتي تسمى الدرجة الخام لا يكون لها معنى، ويصعب تفسيرها ما لم يتم إسنادها الى نظام مرجعي، فهذا النظام هو الذي يسمح باستخلاص معلومات مفيدة من درجات الاختبار. ومن هنا تتضح أهمية معايير الاختبارات، فدرجات اختبارات الذكاء، والشخصية، والاختبارات التحصيلية المقننة وغيرها من الاختبارات والمقاييس تفسر عادة باستخدام جدول المعايير.

المعايير، والمتغيرات، والمستويات

اولاً: المعاييــر-

المعايير عبارة عن جداول تكون ضمن كراسة تعليمات الاختبار تبين بوضوح الدرجات التي حصل عليها المختبر في عينات التقنين، وهي استخدام الإحصاء لوصف أداء المجموعات على الاختبار، وهي درجات مجموعة كبيرة العدد ممثلة للمجتمعات الأصلية التي اشتقت منها.

وهناك صعوبات كبيرة تعترض الأخصائي النفساني في قياسه، أو في تفسيره لخصائص السلوك الإنساني. ولتخطي هذه الصعوبات نشأت الحاجة إلى مفهوم المعايير كون الدرجة الخام التي ينالها الفرد على المقياس لا تفسر شيئاً في تقديرنا لأداء الفرد، إلا بعد مقارنتها بغيرها من الدرجات التي نالها أفراد من نفس

المجموعة التي ينتمي إليها الفرد. إذن لا بد من الرجوع إلى معايير يحدد معنى هذه الدرجة الخام التي نالها الفرد. وتدلنا هذه المعايير على مركز الفرد بالنسبة للمجموعة التي ينتمي إليها.

فالمعيار عادة هو الدرجة المتوسطة عند مجموعة أفراد معينة.

وهذه الدرجات المتوسطة تكون نقطة الإرتكاز التي يتم من خلالها التعرف على مدى تقدم أفراد هذه المجموعة. فالفرد الذي ينال الدرجة الوسط، قد يجعل من أدائه مرجعاً يحدد على أساسه الأداء الضعيف، أو الجيد عند أفراد المجموعة.

والدرجة المعيارية هي المسافة التي تبتعد فيها الدرجة الخام عن المتوسط الحسابي لمجموعة معينة ومعبراً عنها بوحدات من الإنحراف المعياري، وللحصول على الدرجة المعيارية يكون كالتالي:

الدرجة المعيارية = الدرجة الخام – المتوسط ÷ الإنحراف المعياري

لذا فالنتائج التي نحصل عليها من تطبيق الاختبارات والقياسات هي أرقام لامعنى لها، وتدعى درجة خام، وتعرف بأنها النتيجة الأصلية المشتقة من تطبيق الاختبارات أو أي أداة قياس أخرى قبل أن تعالج إحصائيا. وعليه لابد من التعامل إحصائيا مع الدرجة الخام لتحويلها إلى درجة معيارية. والدرجة المعيارية هي درجة يعبر فيها عن درجة كل فرد على أساس عدد وحدات الانحراف المعياري لدرجته عن المتوسط. ويطلق عليها أحيانا اسم (المسطرة) ومن اشهر الدرجات المعيارية (الدرجة الزائية، التائية، المئينية).

تمتاز الدرجة المعيارية بالخصائص الآتية:

1- تحمل معنى واحد من اختبار لآخر، وبذلك يتوفر لدينا أساس للمقارنة بين اختبارات مختلفة.

2- تتألف من وحدات متساوية الأبعاد، بحيث أن الحصول على خمسة نقاط في أحد أجزاء المقياس يكون له دلالة مماثلة للحصول على خمسة نقاط في جزء آخر من المقياس.

3- لها صفر حقيقي يعبر عن (انعدام) الصفة المقاسة، بحيث يصح وصف درجات معينة بأنها تمثل (ضعفي كمية معينة) أو (ثلثي تلك الكمية) وغيرها.

أما أهمية المعايير فتتلخص بالآتي:

- إنها أساس للحكم على الظاهرة من الداخل.

- تأخذ الصيغة الكمية في أغلب الأحوال، فهي تشير لمركز الفرد بالنسبة للمجموعة.

- تتحدد في ضوء الخصائص الواقعية للظاهرة (ما مدى بعد الفرد عن متوسط المجموعة التي ينتمي إليها).

- تعكس المستوى الراهن للفرد.

- وسيلة من وسائل المقارنة والتقويم.

- مهمة في الاختبارات التي تكون على شكل بطارية (فالبطارية هي مجموعة من الاختبارات تعطى لنفس الأشخاص) نظرا لاختلاف وحدات قياس الاختبارات التي تتضمنها البطارية كالثانية، والسنتيمتر، وعدد مرات التكرار...الخ حيث تحول الدرجات الخام (المختلفة بوحداتها) لدرجات معيارية (موحدة في وحداتها) فتسهل بذلك عملية التقويم.

- يمكن الاستفادة منها في التنبؤ وفي تشخيص نواحي القوة والضعف وغيرها.

من متطلبات إعداد المعايير:

1- تحديد الاختبارات المطلوبة، وشروط تطبيقها وتعليمات الإجراء، واحتساب الدرجة وغيرها.

2- اختيار عينة تدعى عينة التقنين [يشترط فيها اعتدالية التوزيع والاختيار بالطريقة العشوائية – كبر حجمها، تمثيلها للمجتمع].

3- تطبيق الاختبارات واستخراج الدرجات المعيارية.

4- مراعاة التوقيت الزمني، فالمعايير دائما مؤقتة لكونها قابلة للتغيير مع مرور الوقت.

5- مراعاة طريقة العرض، إذ يجب أن يتم بجداول واضحة يمكن التعامل معها بسهولة لأغراض التشخيص أو المقارنة ...الخ.

أنواع المعايير:

- ✓ معايير الصف: فيها يقارن الفرد بالمجموعة التي يتساوى أداؤه معها من خلال مجموعات صفية متتابعة.
- ✓ معايير العمر: فيها يقارن الفرد بالمجموعة التي يتساوى أداؤه معها من خلال مجموعات عمرية متتابعة.
- ✓ المعايير المئينية: هي النسبة المئوية من المجموعة التي يتفوق عليها الفرد من خلال مجموعة واحدة عمريه أو صفية ينتمي إليها الفرد.
- ✓ الدرجات المعيارية: هي عدد الانحرافات المعيارية التي يقع عندها الفرد فوق متوسط المجموعة أو تحته من خلال مجموعة عمريه أو صفية ينتمي إليها الفرد.

استخراج المعايير:

- ➢ المرجع المحَكّي : وهو مرجع يندر استخدامه في المقاييس النفسية، وعادةً يقتصر استخدامه في الاختبارات التحصيلية ويعتمد استخدامه على قرارٍ مسبَق بمقدار الأداء المقبول من الفرد،

ولكي نستطيع تصميم اختبار محكّي المرجع فلا بد من وضع أشكال الأداء المطلوبة بصورة واضحة قبل بناء الاختبار، وبناءً عليه فلا بد من تحليل محتوى العمل أو الكتاب الذي نريد أن نقيس من خلاله أداء المفحوص تحليلاً يوضّح العناصر الرئيسة التي يدور حولها الكتاب أو العمل أو أشكال الأداء المطلوبة في

مهارةٍ معيّنة، وبذلك فهذا المعيار معيارٌ قبليٌّ مطلَق، ويكون الحكم على درجات الطلاب هو ما حققه الطالب من أهداف مدروسة وليس بناءً على مقارنته بمجموعته، وفي هذا النوع من المحكّات توضع أسئلة الاختبارات متسلسلة في مستويات الصعوبة التي تمثّل الحد الأدنى المقبول من الأداء والحد الأعلى، ومن أمثلة هذه المحكّات الكتاب المدرسي متمثّلاً في الأهداف التي يراد أن يحققها ذلك الكتاب. ويتم تحديد هذا المحَك قبل تطبيق الاختبار، كما لا يخضع هذا المحك لتقديرات تتعلق بالمستوى العام لأداء أفراد عيّنة معيّنة.

- المرجع المعياري: يختلف هذا المرجع أو هذا المحَكّ عن المرجع المحَكّي اختلافاً جوهريّاً، فنحن نقارن الدرجة التي يحققها الطالب وفقاً لأداء بقية الأفراد، وبذلك فهذا المرجع نهدف من خلاله إلى التعرف إلى الوضع النسبي للفرق بين مجموعة أفراد، فهو معيار بُعْدي أو نسبي، ويعتمد كثيراً على مجموعة من المفاهيم الإحصائية، ويجب أن نتجه في تصميمنا للاختبارات ذات المرجع المعياري إلى أن تكون البنود ذات قدرة على تجسيم مدى الدرجات التي من الممكن أن نحصل عليها عن أي مجموعة من الأفراد.

❖ ثانياً: المتغيرات **Variables**-

مفهوم المتغير (**Variable**):

المتغير هو مقدار له خصائص كمية (رقمية) أو غير كمية (وصفية أو نوعية) تتغير قيمته من عنصر الى آخر من عناصر المجتمع الإحصائي أو العينة. وتُعرف المتغيرات بأنها لا تبقى على وتيرة واحدة. فالأصل فيها الاختلاف والتنوع. وهذا الاختلاف نقصد به تنوع مستويات ذلك المتغير سواء كان ذلك بقيمها الرقمية (السن والدخل والسرعة ومعدل الإنتاج...الخ)، أم بتنوع مستوياتها الاسمية مثل نوع البيئة: (حضر، ريف، بادية، أو نوع الوظيفة: إدارية، كتابية، فنية أم الجنسية :سعودي، مصري، يمني...الخ) كما يشترط فيها قابليتها للقياس ((Bouma,1993)). والحد الأدنى من المستويات هو اثنان بالنسبة للمتغيرات الاسمية، وثلاثة مستويات أو أكثر للمتغيرات الأخرى (المتغيرات الرتبية والفئوية والرقمية والنسبية). فالمتغير شيء ومستوى المتغير شيء آخر :

فعندما نقول مثلاً الجنس (النوع) فالحديث هنا عن متغير وصفي لحالة الكائن الذي نصفه، في حين نقول ذكر أو أنثى فنعني مستويات ذلك المتغير (الجنس) والذي فيه مستويات تختلف فيما بينها، فكل مستوى يمثل خاصية واحدة فقط.

أي أن المستوى يمثل خاصية وصفية واحدة لا تقبل احتمالية أي تكرار لوصف آخر.

أنواع المتغيرات:

تنقسم المتغيرات الى نوعين:

اولاً- المتغيرات المصنفة أو النوعية (**Categorical Variables**):

وهي تلك المتغيرات التي تصنف مفردات (أو عناصر) المجتمع الاحصائي أو العينة في عدة مجموعات، تشترك كل مجموعة في صفة معينة. ومن امثلة المتغيرات النوعية أو الوصفية : متغير نوع العمل الذي يصنف المفردات إلى اربع مجموعات هي (منتج، موظف، عمل حر، أو متقاعد)، ومتغير الجنس الذي يصنف مفردات المجتمع أو العينة إلى مجموعتين هما (ذكر، انثي)، ومتغير الحالة الاجتماعية (أعزب، متزوج) هو، كذلك متغير مستوى الدخل الذي يمكن تصنيفه الى: عالي، متوسط، منخفض.

ثانياً- المتغيرات الكمية (**Quantitative Variables**):

وهي المتغيرات التي تأخذ قيماً كمية (أو عددية) .. لذا فانه يطلق عليها اسم المتغيرات الكمية. ومن أمثلة المتغيرات الكمية: أعداد الطلبة في كل قسم علمي بكلية العلوم في جامعة القدس، والوزن (بالكيلو جرام) لعينة من الأطفال الدارسين بالصف الأول الابتدائي في مدينة معينة، والعمر (بالسنوات) للموظفين العاملين بإحدى الشركات النفطية. وتنقسم المتغيرات الكمية الى قسمين هما:

1- المتغيرات المنفصلة أو المتقطعة (Discrete Variables):

وهي المتغيرات الكمية التي تأخذ قيما عددية محددة صحيحة، ولا اتحتوي علي قيم كسرية، مثل: عدد المصانع في كل مدينة من مدن دولة ما، عدد حوادث السيارات التي وقعت في الربع الأول من هذه السنة في 5 مناطق مختلفة، وعدد المدرسين في كل مدرسة من المدارس الثانوية في مدينة رام الله خلال هذه السنة، ....الخ.

2- المتغيرات الكمية المتصلة (Continuous Variables):

وهي المتغيرات الكمية التي تأخذ قيما تكون عددا صحيحاً وكسرا من وحدة القياس مثل : متغير الدخل اليومي (بالدينار) لعينة من الأشخاص، الطول (مقاسا بالسنتمترات)، الوزن (مقاسا بالكيلو جرام)، العمر (مقاسا بالسنوات) درجة حرارة الجو (مقاسة بالدرجات)، المساحة (مقاسة بالمتر المربع) وغيرها.

أقسام المتغيرات النوعية:

هذا وتقسم المتغيرات النوعية أو المصنفة حسب المقياس الذي قيست به إلى قسمين هما :

أ- متغير إسمي (Nominal Variable): يقال بإن المتغير النوعي متغير إسمي عندما لا تخضع صفاته لاي معيار من معاير الترتيب. وتكون الصفات أو الانواع في المتغيرات الإسمية عبارة عن اسماء. ومن أمثلة المتغيرات الاسمية: الديانة: التي يمكن تصنيفها إلى [مسلم، مسيحي، يهودي...] الجنس: الذي يمكن تقسيمه إلى [ذكر، انثي].

ب- متغير ترتيبي (Ordinal Variable): يقال بإن المتغير النوعي متغير ترتيبي إذا امكن وضع صفاته أو اقسامه في ترتيب تصاعدي (أو تنازلي). ومن أمثلة المتغيرات الترتيبية : المستوي التعليمي، والذي يمكن تصنيفه الى (عالي, متوسط , منخفض)، التقديرات المختلفة للنجاح في الثانوية العامة (مقبول , جيد , جيد جدا , ممتاز)، تأييد قضية معينة، والذي يمكن تصنيفه الى [اوفق , متردد , لا اوافق].

كما تقسم المتغيرات الكمية (المتصلة أو المنفصلة) أيضا حسب المقياس الذي قيست به إلى قسمين هما :

1- متغير فئوي (Interval Variable):

يقال بإن المتغير الكمي فئوى عندما تكون المسافات بين النقاط في المقياس الذي قيس به المتغير متساوية، حيث يمكن في المتغيرات الفئوية مقارنة بين القيم، ولا يمكن المقارنة بين نسب القيم. ومن أمثلة المتغيرات الفئوية درجة الحرارة المئوية، حيث مقاييس درجة الحرارة المئوية تؤلف أو تكون متغير كمي فئوي. فمثلاً اذا كان لدينا مفردات درجة حرارتها هي: 12 ـ 24 ـ 36 على التوالي، فإنه من المناسب أن نقول بان الفرق بين درجة حرارة المفردة الاولي والثانية يساوي الفرق بين درجة حرارة المفردة الثانية والثالثة. كما انة من المناسب ان نقول بأن الفرق بين درجة حرارة المفردة الاولى والثالثة هو ضعف الفرق بين درجة حرارة المفردة الاولى، والثانية أو المفردة الثانية والثالثة. ولكن ليس من المناسب ان نقول

بان درجة حرارة المفردة الثانية تساوي ضعف درجة حرارة المفردة الاولى، أو أن نقول بـإن درجة حرارة المفردة الثالثة هي ثلاث أمثال درجة حرارة المفردة الاولى.

2- متغير نسبي (Ratio Variable):

يقال بأن المتغير الكمي متغير نسبي، عندما يمكن المقارنة بين نسب القيم الممكنة للمتغـير بالاضافة الى مقارنة الفرق بين القيم، حيث يمكن في المتغيرات النسبية قسمة قيمها الرقمية بحيث تعطي نسـباً ذات معنـى. فـالطول والـوزن والـدخل مـن أمثلـة المتغـيرات الكميـة النسبية.

وعلى الرغم من تنوع المتغيرات بناءً على تقسيماتها، إلا أنها من حيث القيـاس تنقسـم إلى ثلاث أنواع على النحو التالي:

نوع المتغير	نوع القياس
المتغير الاسمي	قياس اسمي
المتغير الرتبي	قياس رتبي
المتغير المنفصل، والمتصل، والفئوي، والنسبي	قياس كمي

وعلى الرغم التصنيفات السـابقة للمتغـيرات إلا إنهـا تتميـز بخاصـية التحـول مـن شكل إلى آخر، أي أنه يمكـن تحويـل متغـير إلى متغـير آخـر، ويمكننـا تشـبيه تلـك العملية بعملية نقل الدم بين الفصائل. حيث إن المتغير الاسمي يقبل أن تحول لـه

جميع المتغيرات (المستقبل العام) والمتغير الرقمي يمكن أن يتحول إلى أي نوع من المتغيرات (المعطي العام) لكن لا يمكن التحويل إليه.

ويعتبر المتغير الاسمي المستقبل العام، أي أنه المتغير الذي يمكن أن تتحول له جميع المتغيرات والعكس غير صحيح.

ولتوضيح ذلك نضع المثال التالي:

1- تحويل متغير كمي (رقمي، متصل، فئوي، نسبي) إلى متغير اسمي:

لو أراد باحث التعامل مع متغير رقمي يمثل العمر مثلاً على اعتبار أنه متغير رقمي فإن عليه أن يحول مستويات ذلك المتغير إلى مستويين فقط (كبير وصغير) فهنا يقيس الاختلافات بين تلك الفئتين فقط. في هذه الحالة نحتاج إلى عينتين الأولى لفئة كبار السن والأخرى لفئة صغار السن.

2- تحويل متغير رتبي إلى متغير اسمي:

لو أن باحث يريد تحول فئات متغير التعليم الرتبية (ابتدائي، متوسط، ثانوي، جامعي، عالي) إلى متغير اسمي، فتتسم العملية على النحو التالي: تعليم منخفض للمستويات (ثانوية فأقل)، وتعليم عال لمن تعليمهم فوق الثانوية.

ويقبل المتغير الرتبي تحويل المتغيرات كلّها عدا المتغير الاسمي، على النحو التالي:

- تحويل متغير كمي (رقمي، متصل، فئوي، نسبي) إلى متغير رتبـي: لـو رجعنـا إلى المتغـير الفئوي والخاص بالعمر والذي كان له الفئات التالية :

(من 18 إلى أقل من 32)

(من 32 إلى أقل من 46)

(من 46 إلى أقل من 60)

يمكننا أن نحول ذلك المتغير إلى الفئوي إلى متغير رتبي على النحو التالي

(من 18 إلى أقل من 32) = صغير

(من 32 إلى أقل من 46) = متوسط

(من 46 إلى أقل من 60) = كبير

وبذلك أصبح متغير العمر الفئوي متغير رتبي.

وتساهم عملية تحويل المتغيرات فيما بينها إلى القدرة على تطبيق قياسات متنوعـة تمكـن الباحث التعرف وبشكل أعمـق عـلى طبيعـة العلاقـات بـين المتغـيرات، كـما تسـهم هـذه العملية في التغلب على بعض العقبات التي قد تنشأ نتيجة لبعض خصائص العينـة كـأن لا تمثل فئات معينة في مفردات الدراسة.

موقع المتغير بالنسبة إلى الدور الذي يلعبه:

أما من حيث موقع المتغير بالنسبة إلى الدور الذي يلعبه في ارتباطه في علاقته مع المتغيرات الأخرى فيمكن أن نقسمه إلى ثلاثة أنواع رئيسية هي:

- المتغير المستقل Independent Variable
- المتغير التابع Dependent Variable
- المتغير الثالث (الدخيل) Variable extraneous

☒ المتغير المستقل **Independent Variable:**

يُعرَّف المتغير المستقل بأنه المتغير الذي يحدث قبل حدوث المتغير التابع وذلك من الناحية الزمنية. كما أنه لا يتأثر بأية متغيرات أخرى، فالاستقلالية تعني أنه غير مرتبط بمتغير آخر. وقد تشترك مجموعة من المتغيرات المستقلة في التأثير على المتغير التابع. ويعد ذلك من الأمور الكثيرة التكرار في الدراسات المسحية، أو أن يكون المتغير المستقل سبباً في حدوث المتغير التابع وذلك في الدراسات التجريبية لمعرفة تأثير ذلك المتغير المستقل على المتغير التابع، ويتم ذلك بعد التحكم في جميع المتغيرات المستقلة الأخرى التي يمكن أن يكون لها تأثير على المتغير التابع.

ويُعرف المتغير المستقل في هذه الحالة بالمؤثر.

☒ المتغير التابع **Dependent Variable:**

المتغير التابع هو ذلك المتغير الذي ظهر أو تشكل نتيجةً لتدخل متغير أو متغيرات أخرى كالمتغير أو المتغيرات المستقلة والتي ساهمت في ظهوره بهذه الحالة. وتعد معرفة ذلك التأثير الحادث على المتغير التابع من الأساسيات لفهم العلاقة التي يريد الباحث الوصول إليها وتفسيرها واستنتاجها من خلال دراسته.كما أن المتغير التابع في دراسة ما لا يبقى على وتيرة واحدة فقد يدرسه الباحث نفسه أو باحث آخر على أنه متغير مستقل في دراسة أخرى (Royes, 1991).

فعلى سبيل المثال لو أراد باحث دراسة ارتباط الطلاق بانحراف الأبناء. فهنا يعد المتغير التابع هو انحراف الأبناء. ولو فرضنا أن باحثاً يريد دراسة تأثير انحراف الأبناء بمعدل الطلاق، في هذه الحالة انحراف الأبناء هو المتغير المستقل.

☒ المتغير الثالث (الدخيل): **Variable extranious**

هومتغير لا يرغب الباحث في دراسته لكنه قد يشوش على المتغير المستقل و ينافسه في التأثير على المتغير التابع.

ثالثاً: مستويات القياس-

القياس في معناه الدقيق ما هو إلا تعيين أو تحديد أعداد للظواهر المقاسة أو التي نلاحظها، وذلك بالطريقة التي تيسر لنا تحليل تلك الأعداد وفقاً لقواعد معينة،

ولكن هل معنى هذا أن كل الظواهر تخضع لقواعد الأعداد المعرفة ؟ (قابلية الجمع، تساوي الفروق بين الأعداد المتتالية، قابلية الترتيب......)، في حقيقة الأمر أن القواعد المعروفة للأعداد لا تنطبق على كل الظواهر المقاسة، وهو ما يقودنا إلى التفرقة بين المستويات المختلفة للقياس:

1- المستوى الاسمي (التصنيفي):

يعد هذا المستوى أدنى مستويات القياس وأضعفها، بل أن تسميته مقياساً تعد أحياناً تسمية مجازية، وهنا تستخدم الأعداد بغرض التصنيف فقط لا غير مثال: تحديد العدد 1 ليدل على أن جنس المفحوص ذكر والعدد 2 ليدل على أن الجنس أنثى، فهنا لا يعني ذلك أن 2 أكبر من واحد وإنما الغرض تصنيفي بحت وهكذا تحديد بعض الأعداد لتدل على التخصص، ولون العين والجنسية كلها من نوع المقاييس الاسمية أو التصنيفية، وبالتالي فالقياس الاسمي أو التصنيفي يعني بتصنيف الأفراد في الظاهرة، ولكن لا يوضح ترتيبهم في هذه الظاهرة.

2- المستوى الترتيبي (الرتبي):

وهنا الهدف من الأعداد هو ترتيب الأفراد في ظاهرة معينة، ولكن يجب ملاحظة أن الفروق بين الرتب ليست متساوية أو لا يعني أن المسافات البينية بين الأفراد متساوية، فمثلاً: إذا حددنا الأعداد من واحد إلى خمسة لتدل على ترتيب الأفراد في ظاهرة القلق أو النشاط الحركي الزائد فلا يكون الفرق بين التلميذ الأول (الأقل في القلق) والتالي له هو نفس الفرق بين التلميذ الخامس (الأكثر) في

القلق والسابق له، ويجب ملاحظة أن القياس الترتيبي قد يحمل معنى التساوي أحياناً، فالقياس الترتيبي يعطي فكرة عن ترتيب الأفراد في ظاهرة معينة ولكن لا يعطى فكرة عن الفروق في الظاهرة بين الأفراد.

3- المستوى الفتري أو الفئوي:

وهو هنا تتساوى الفروق أو المسافات بين المستويات المتتالية مثل درجات الأطفال في اختبار الذكاء أو درجاتهم في التحصيل أو درجاتهم في اختبار القلق، فهنا يكون وحدة قياس ثابتة متفق عليها يقاس بها الفرق بين كل درجة والتالية لها بحيث يصبح الفرق مثلاً بين 4- 5 مساوي للفرق بين 10- 11 أو المسافة بين 10- 15 مساوي للفرق بين 20- 25 ولا يعنى ذلك وجود صفر مطلق، يعنى غياب الصفة، وإنما البداية أو الصفر هنا صفر اختياري أو نسبي وليس صفراً مطلقاً ويجوز إجراء العمليات الحسابية لتقليدية كالجمع والطرح.

4- المستوى النسبي:

وهنا للدرجات صفر مطلق يعنى غياب الصفة مثل مقياس الوزن ودرجة الحرارة، ويسمى هذا المستوى بالمستوى النسبي، لأن النسبة بين أي درجتين لا تتأثر بوحدة القياس، فمثلاً النسبة بين واحد كيلو جرام و10 كيلو جرام هي نفسها النسبة بين 1000 جرام و 10.000 جرام، وهنا الأعداد المستخدمة أعداد حقيقة لها صفر مطلق.

ولا نتطلع في العلوم الإنسانية دائماً أن نصل لأعلى أو أكثر من المستوى الفتري أو الفئوي، فلا يمكننا القول بأن الطفل الذي حصل على نسبة ذكاء 140 يساوى في ذكاءه ضعف الطفل الذي حصل على نسبة ذكاء 70، ولكن إذا كنا بصدد التعرف على العلاقة بين الوزن أو الطول وأي من المتغيرات النفسية، فإن الوزن أو الطول في هذه الحالة يقاس في المستوى النسبي.

أهمية المستويات:

- ✓ أسس داخلية للحكم على الظاهرة.
- ✓ تأخذ الصورة الكيفية.
- ✓ تتحدد في ضوء ما يجب أن تكون عليه الظاهرة.
- ✓ يتم أعدادها على أفراد مدربين ذوي مستويات مثالية، كما يتم أعدادها بعد التعلم والتدريب والممارسة بهدف التحصيل أو تطوير الصفة أو الخاصية للوصول لدرجات تعكس المستوى الأمثل للصفة أو الخاصية.

تستخدم المستويات في تقويم المستوى من خلال المقارنة بمحك، أو في تقدير مستويات الإنجاز في الأنشطة التي تتطلب الإتقان (أي مستوى نطاق المحتوى)، حيث يستخدم في اختبارات التحصيل لتفسير الأداء من خلال ملاحظة ما يؤديه الفرد فعليا بالمقارنة بما يجب أن يكون عليه الأداء، وليس مقارنة بأداء الآخرين (أي الحكم هنا على مدى الإتقان)، أو استخدام المعايير الارتقائية الرتبة تلك التي تعتمد على الوصف الكيفي للسلوك الذي يجب أن يكون عليه الفرد، مثل

اعتماد نتائج بحوث علم نفس النمو التي وصفت السلوك الإنساني في المراحل المتتابعة (كجداول جيزل الارتقائية بالولايات المتحدة الامريكية).

أخلاقيات استخدام الاختبارات:

بعض ضوابط وأخلاقيات استخدام الاختبارات والمقاييس النفسية والتربوية:

يتضح مما سبق عرضه من شروط علمية واعتبارات فنية لمختلف مراحل العملية الاختبارية أن استخدام الاختبارات النفسية علم وفن، فالاختبارات هي أدوات قياس تعين الاخصائيين في مختلف مجالات الخدمات النفسية في أداء واجباتهم الإنسانية بأسلوب علمي موضوعي واستخدامها في تحقيق الأهداف التي أعدت من أجلها يتوقف على نوعية هذا الاستخدام، وكفاءة وخبرة العاملين في هذا المجال.

ولضمان التزام جميع الاطراف المعنية باستخدام الاختبارات والمقاييس النفسية، ينبغى أن تكون هناك مجموعة من الضوابط والأخلاقيات التي توجه المسار، وتحدد الممارسات في هذا المجال، وتضفي على المهنة الصفة الكاملة التي تستوجبها من المجتمع، وتراعي حقوق وواجبات القائمين بهذه الخدمات والمستفيدين منها.

ولذلك حرصت الجمعية الأمريكية لعلم النفس (APA) على إصدار دليل متجدد يتضمن الاسس والمبادئ الاخلاقية التي يجب أن يلتزم بها الأخصائيون

النفسيون في أدائهم لمسؤولياتهم، ومن بينها المبادئ الأخلاقية المتعلقة باستخدام الاختبارات النفسية وأدوات القياس التشخيصية.

كما أصدرت دليلاً للمعايير التي ينبغى توافرها في هذه الاختبارات، وإرشادات عن إجراءات استخدامها. وكذلك أصدرت الجمعية الأمريكية لشؤون الأفراد والتوجيه دليلاً يتضمن المبادئ الأخلاقية في مجالات الخدمات النفسية متضمنة مجال تطبيق الاختبارات والمقاييس.

وعلى الرغم من أن هذه الضوابط والأخلاقيات يجب أن تنبع من واقع أعضاء المهنة أنفسهم، إلا أننا سوف نقدم بعض التوجيهات التي يمكن الاسترشاد بها في هذا الشأن.

الكفايات اللازم توفرها فيمن يستخدم الاختبارات: إن وضع ضوابط للكفايات اللازم توافرها فيمن يستخدم الاختبارات والمقاييس النفسية تعد خطوة أساسية لتجنب إساءة هذه الاستخدامات، وتختلف بالطبع هذه الكفايات باختلاف نوع الاختبار والهدف منه، ويمكن تصنيف الاختبارات والأدوات التشخيصية والكفايات اللازم توافرها فيمن يستخدمها إلى ثلاثة تصنيفات (APA،1973):

التصنيف الأول: اختبارات أو أدوات تشخيصية يمكن تطبيقها وتصحيحها وتفسير نتائجها بأسلوب مناسب بالرجوع إلى كراسة تعليمات الاختبار ومراعاة مطالب المؤسسة التي يعمل فيها مستخدم الاختبار، ومن أمثلة ذلك الاختبارات التحصيلية، واختبارات الكفايات المهنية، ويمكن أن يقوم بتطبيق هذه الاختبارات

وتفسير نتائجها معلمين ومسؤولين عن المؤسسات المختلفة بحيث يكونـون ملمـين بدرجـة معقولة بأساسيات القياس التربوي والنفسي.

التصنيف الثـاني: اختبـارات وأدوات تشخيصـية يتطلـب اسـتخدامها معرفـة أسـاليب بنـاء الاختبـارات وطـرق اسـتخدامها، ودراسـة بعـض المقـررات في الاحصـاء الـنفسي، والفـروق الفردية، وسيكولوجية التوافق، وسيكولوجية انتقاء الأفراد، والتوجيه المهني والتربوي. ومن أمثلة ذلك اختبارات الذكاء العام، والاستعدادات الخاصة والاستعدادات الفارقة، ومقـاييس الميول، واستبيانات الشخصية، وينبغي فيمن يستخدم هذه الاختبارات والأدوات أن يحقق هذه المتطلبات المذكورة، أو أن يكون مصرحاً له باسـتخدامها في إطـار مؤسسـة تربويـة أو صناعية أو إدارة حكومية يعمل فيهـا، أو أن يكـون قـد أدى بنجـاح دورة تدريبيـه تتعلـق بهذه الاختبارات.

التصنيف الثالث: اختبارات وأدوات تشخيصية تتطلـب معرفـة متعمقـة في القيـاس النفسي، وبعض المجالات النفسية المصاحبة، وكذلك تدريباً مكثفاً وخبرة عملية تحـت إشراف أخصـائي إكلينـيكي لمـدة لا تقـل عـن عـام. ويفضـل أن يكـون الأفـراد الـذين يستخدمون هذه الأدوات من الحاصلين على درجة الماجسـتير في علـم الـنفس، ومـن أمثلة هذه الأدوات: اختبارات الذكاء الفردية، والأساليب الإسقاطية، وبعـض مقـاييس الشخصية الأخرى.

ويحب على من يستخدم الاختبارات والمقاييس النفسية بعامة أن يكون مـزوداً بالمهـارات التـي تمكنـه مـن انتقـاء الاختبـارات التـي تتفـق والهـدف مـن عمليـة القيـاس،

وتناسب الفرد الذى تجري عليه هذه الاختبارات، وأن يكون مطلعاً على الدراسات والبحوث المتعلقة بالاختبار أو الأداة المستخدمة، وكذلك يمكنه تقييم الخصائص السيكومترية للاختبارات، من مثل المعايير، والثبات، والصدق، وأن يكون على دراية بالعوامل النفسية والظروف المختلفة التي تؤثر في العملية الاختبارية وفي الأداء، ويمكنه التوصل إلى استنتاجات وتوصيات اعتماداً على درجة أو درجات الاختبار، آخذاً بعين الاعتبار المعلومات الأخرى المهمة المتعلقة بالمفحوص.

حقوق المخَتَبرين:

إن من حق الفرد الذى يجرى عليه اختبار أو مقياس معين أن يكون على علم بالهدف من الاختبار وفيم تستخدم نتائجه. فلا يجوز للأخصائي النفسى أن يطلب من المفحوص الإفصاح عن معلومات تخصه أثناء عملية المقابلة الشخصية أو أثناء اختباره أو تقويمه، أو أن يسعى إلى الحصول على مثل هذه المعلومات إلا بعد تعريفه بالهدف من ذلك وأخذ موافقته كتابياً، أو أخذ موافقة ولى أمره إذا كان قاصراً.

عدم تداول الاختبارات بين غير المختصين في علم النفس: إن الحفاظ على سرية الاختبارات تعد من مسؤولية مؤلفيها وناشريها ويشاركهم في ذلك الاخصائي النفسي (Sax, 1974) فسرية الاختبارات تعد مشكلة إذا أدت إلى تغيير في الدرجة التي يحصل عليها الفرد في الاختبارات دون أن يكون هناك تغيير مقابل في درجته الحقيقية، غير أن سرية الاختبارات لا تعد مشكلة في بعض

الاختبارات. فمثلاً اذا أُجري اختبار على فرد لقياس درجة تمكنه من مهارة معينة، فإن معرفته السابقة بعينة من أمثلة مشابهة لأسئلة أو تدربه عليها ربما يوصى به.

ولكن في كثير من الحالات تؤثر معرفة عينة من الأسئلة أو طريقة تصحيحها تأثيراً بالغاً في صدق الاختبار، لذلك يجب في مثل هذه الحالات الحفاظ على الاختبارات في زوايا مغلقة، وتحديد الأشخاص الذين توكل إليهم مهمة استلام هذه الاختبارات.

وإذا أُرسلت الاختبارات إلى مكان آخر عن طريق البريد، فإن الشخص المرسل إليه يجب أن يكون معروفاً وموثوقاً به. والواقع أن آلات التصوير الحديثة والتقنيات المختلفة قد زادت من حدة مشكلة سرية الاختبارات وما يرتبط بها من مخالفات لحقوق الطبع والنشر، مما يتطلب اتخاذ الإجراءات التي تضمن منع هذه المخالفات.

سرية تقارير نتائج الاختبارات:

إن من حق المفحوص الاطلاع على تقرير نتائج الاختبارات التي أداها، كما يجب أن يسمح له بالتعليق على محتوى التقرير، وتوضيح أو تعديل بعض المعلومات الخاصة به عند الضرورة. ويهتم المرشدون النفسيون في الوقت الحاضر بأن يشارك المفحوص مشاركة فاعلة في عملية تقويمه، لذلك يجب أن تقدم إليه نتائج الاختبارات بشكل يسهل عليه فهمه، وأن يبتعد التقرير عن المصطلحات والعبارات الفنية المعقدة.

واذا كان المفحوص طفلاً يكون من حق والديه الاطلاع على تقرير النتائج، وإن كان هذا يمثل تدخلاً في خصوصياته Ruebhausen،1961، غير أن التوجيهات التي اقترحتها مؤسسة زيجRussell Sage Foundation،1970 تشير إلى أنه إذا بلغ الطفل من العمر ثمانية عشر عاماً، ولم يتمكن من استكمال دراسته الثانوية، أو تزوج، فإن من حقه أن لا يطلع على مثل هذا التقرير سواه.

غير أن هذا التوجيه أعقبه تحذير حول مدى إمكانية تنفيذ ذلك، ومراعاة القوانين المتعلقة بهذا الشأن. ولكن المشكلة هنا ليست مجرد اطلاع القاصر أو والديه على تقرير نتائجه، وإنما كيفية إجراء ذلك. فمن الناحية القانونية للوالدين الحق في الحصول على معلومات تتعلق بأبنائهم، وهم عادة يحرصون على ذلك، إلا أنه في بعض الحالات ربما تزداد حدة المشكلات التعليمية والانفعالية للطفل نتيجة عدم توافقه مع والديه. ففي مثل هذه الحالات ينبغى تدخل المرشد النفسي لكي يحصل على المعلومات اللازمة عن الطفل من والديه والعمل على كسب تعاونهم. أما إذا طبقت الاختبارات من خلال مؤسسات عمل أو مؤسسات تربوية أو قضائية فإنه ليس من الضروري أخذ موافقة كتابية من الفرد المختبر، وإنما يلزم إشعاره بموعد الاختبار والهدف منه وكيف ستستخدم نتائجه، ومن هم الذين سوف يستخدمونها. والمشكلة المهمة الأخرى تتعلق بإستبقاء المعلومات المتعلقة بنتائج الاختبارات النفسية والاحتفاظ بها، فهذه المعلومات تفيد في الدراسات الطولية، وتتبع حالة المختبرين. غير أن هذه المعلومات تتقادم بعد مدة زمنية تختلف باختلاف نوعيتها، وبذلك تؤدي إلى استنتاجات غير واقعية،

واستخدامها في غير الأغراض التي جمعت من أجلها. لذلك يجب أن تكون هناك ضوابط تنظم عمليات استبقاء هذه المعلومات والحفاظ عليها وعدم النيل منها أو الوصول اليها إلا في الحالات المسموح بها.

ومما يزيد هذه المشكلة تعقيداً التطورات الكبيرة التي حدثت في مجال نظم المعلومات وإمكانية استدعاء المعلومات في سرعة ودقة متناهية، إلا أنه يمكن الإفادة من هذه التطورات في الحفاظ على المعلومات والبيانات المتعلقة بنتائج الاختبارات وابتكار أساليب تضمن سرية هذه النتائج.

نشر وتوزيع الاختبارات:

يمكن تلافي كثير من المشكلات الناجمة عن سوء استخدام الاختبارات بتنظيم عملية نشر وتوزيع الاختبارات، فبعض الناشرين يسمحون ببيع الاختبارات والمقاييس النفسية لكل من يطلبها مادام أنه يدفع ثمن شرائها، في حين أن عدداً محدوداً منهم يقصر البيع على فئة من المتخصصين، وربما يرجع السبب في ذلك إلى أن بعض هؤلاء الناشرين لا يدركون الأضرار التى تنجم عن سوء توزيع الاختبارات سواء للأفراد أو للمؤسسات.

لذلك فإن هناك حاجة إلى تقييد توزيع أو بيع الاختبارات والمقاييس النفسية، وقصر ذلك على من تتوفر لديهم المهارات أو الكفايات اللازمة لاستخدامها الاستخدام المناسب، وذلك لضمان سرية الاختبارات وعدم إساءة استخدامها.

ومن مسؤولية القائمين بإعداد الاختبارات أو تأليفها، ومؤسسات نشرها عدم السماح بتوزيع الاختبارات للاستخدام قبل أن يتم إعدادها في صورتها النهائية. وإذا سمح بنشر اختبار في صورته الأولية لأغراض البحث أو الدراسة يجب أن يوضح ذلك صراحة على غلاف الاختبار وفي كراسة تعليماته، وأن يقتصر توزيعه على تلك الأغراض فقط. ويجب أن تشتمل كراسة تعليمات الاختبار بشكل عام على بيانات كافية يمكن الاستعانة بها في تقييم الاختبار.

وكذلك تشتمل على معلومات عن كيفية تطبق الاختبار وتصحيحه وتفسير درجاته في ضوء المعايير الخاصة به، وأن تكون الإرشادات والمعلومات الواردة تعبر حقيقة عما يهدف له الاختبار.

كما يجب على مؤلفي الاختبارات أن يعيدوا النظر في اختباراتهم ومعاييرها كلما دعت الحاجة إلى ذلك حتى لا تتقادم هذه المعايير، مع مراعاة أن الاختبارات تختلف اختلافاً ملحوظاً في المدة الزمنية اللازمة لمراجعتها.

ومن الضوابط أيضاً في مجال نشر الاختبارات عدم السماح بالإعلان عن الاختبارات أو نشر أجزاء منها في الصحف أو المجلات أو الكتب العامة حتى لا يساء فهمها أو استخدامها، وأن يقتصر النشر على الدوريات العلمية المتخصصة بأسلوب موضوعي بعيد عن الانفعال أو الإقناع أو الوصف.

الفصل الخامس

مصطلحات شائعة في القياس النفسي:

- اختبار استعداد: (Readiness Test) اختبار يُحدد درجة امتلاك الطالب للمهارات القبلية الضرورية للنجاح في نشاط تعلم معين.
- اختبار الجاهزية: (Readness Test) هو مقياس لدرجة النضج التي وصل اليها الفرد، أو هو مقياس لما اكتسبه من مهارات أو معلومات يحتاجها لكي ينفذ بعض المهام التعليمية بنجاح.
- اختبار السرعة: (Speed Test) اختبار يقاس الأداء فيه بعدد المهام التي يمكن إنهاؤها في وقت محدد. (اختبار سرعة الطباعة، مثلا)، إنه اختبار محدد بوقت قصير بحيث أن عددا كبيرا من المفحوصين لا يستطيعون الإجابة على كل مفرداته في الوقت المتاح.
- اختبار القدرة: (Ability Test) اختبار يصمم ويبنى ليقيس إمكانات الطالب التحصيلية والمهارية الحالية، أو ليتنبأ بقدرته المستقبلية في أحد المجالات المعرفية (الرياضيات، مثلا). وغالبا ما يتم توظيف نتيجة هذا الاختبار لأغراض مثل: تسكين الطالب (placement)، أو تقديم برنامج علاجي له (Remedial program).
- اختبار القوة: (Power Test) هو اختبار لا يعد الوقت فيه عاملا مؤثرا على مستوى أداء الطالب في الاختبار، ويتم ترتيب أسئلته تصاعديا حسب صعوبتها. أي أن الطالب في هذا الاختبار لا يقدم أداء أفضل فيما لو أعطي وقتا إضافيا. يقابل هذا الاختبار اختبار السرعة الذي يعتبر الوقت فيه

عنصرا جوهريا في تقييم الأداء. ويتم تصحيح هذا الاختبار بناء على درجة صعوبة بنوده.

- الاختبار المقنن: Standardized test هو ذلك الاختبار الذي تتوحد فيه وتتحدد بدقة طريقة تطبيقه، ومواده، وتعليمات وزمن إجابته، وطريقة تصحيحه أو تسجيل درجاته، بحيث يصبح الموقف الاختباري موحداً قدر الإمكان لجميع الأفراد. وهذا يجعل من الممكن المقارنة بين درجات الأفراد الذين طبق عليهم الاختبار.
- اختبار بعدي: (Post Test) اختبار أو تقويم ينفذ بعد انتهاء فترة التجربة أو المعالجة (Treatment) أو فترة التدريس بغرض قياس أثر التجربة أو فاعلية طريقة التدريس.
- اختبار قبلي: (PreTest) اختبار يعطي للطلاب قبل بدء التدريس لتقرير مستوى أداء الطالب الراهن في موضوع أو موضوعات معينة.وفي البحث التربوي هو اختبار يعطي لمجموعات الدراسة قبل تطبيق المعالجة (Treatment)، وذلك لتحديد أثر تلك المعالجة (المتغير التجريبي).
- الاختبار كمؤشر: (Sign Test) وهو الاختبار الذي يتم تكوين فقراته بطريقة لا يمكننا القول بأنها تمثل المجال موضع الاهتمام، إذ إن المجال في هذه الحالة لا يمكن تحديده تماماً، ويعرف هذا بأنه مفتوح النهاية. ويشمل ذلك مجال قياس القدرات العقلية كالذكاء وبعض السمات النفسية.

- الاختبار محكي المرجع: (Criterion – Referenced Test) هو أن يقارن أداء الفـرد بمستوى أداء معين يتم تحديده بصرف النظر عن أداء المجموعـة، ويهـدف الاختبـار محكي المرجع إلى تقدير أداء الفرد بالنسـبة إلى المحـك أو مسـتوى أداء مطلـق، دون اللجوء إلى مقارنة أدائه بأداء الأفراد الآخرين.
- الاختبار معياري المرجـع: ((Norm – Referenced Test اختبـار مرجعـي المعيار ذلك الاختبار الذي يستخدم لتقدير أداء الفرد بالنسـبة لأداء الأفـراد الآخرين في القدرة التي يقيسها ذلك الاختبار. ويسمى الدرجة التـى حصـل عليها الدارس درجة معيارية بينما يطلق على المجموعة التي نال فيها هذه الدرجة فئة معيارية.
- اختبار: أداة تقويم يتم إعدادها وفـق طـرق منهجيـه منظمـة، يـتم إعـداد هذه الأداة لغرض تحديـد درجـة إمـتلاك الطالـب لسـمة أو قـدرة معينـة، وذلك من خلال إجاباته عـلى عينـة مـن المثـيرات ذات العلاقـة بالسـمة أو القدرة المرغوب قياسها.
- الاختبـارات الجمعيـة: (Grouped Tests) هـي مجموعـة الاختبـارات التـي يـتم تطبيقها على مجموعة من الأفراد، كما في الاختبارات التحصيلية الصفية والاختبارات المقننة واختبارات القدرات العقلية كما في اختبار كاليفورنيا للنضج العقلي، واختبـار بيتا للذكاء غير اللفظي.

- ❖ الاختبارات الفردية: (Individual Tests) هي مجموعة الاختبارات التي يمكن أن تطبق بصورة فردية (أي تطبيق على كل مفحوص على حده). مثل اختبارات الذكاء ستانفورد بينيه ووكسلر، ويشيع استخدام مثل تلك الأدوات لأغراض إكلينيكية.
- ❖ الاختيار من متعدد: (Multiple Choice Test) اختبار يرد لكل سؤال فيه بديلان أو أكثر من الإجابات البديلة، وعلى الطالب أن يختار الإجابة الصحيحة من بين تلك الإجابات البديلة، علما بأنه لا يوجد سوى إجابة واحدة صحيحة لكل سؤال. تسمى الإجابات غير الصحيحة بالمشتات(distracters)، ويسمى رأس السؤال بالجذر.
- ❖ الأدوات الاسقاطية: هي مجموعة الادوات التي لا يكون أي من المثير أو الاستجابة محددا (غامض) كاختبار البقع الحبرية (الرورشاخ)، واختبار تفهم الموضوع (TAT)، وبعض الاختبارات النفسية الأخرى التي تستخدم في استقصاء نوعية وحجم المشكلات التي يواجهها الأفراد في دراستهم أو في حياتهم الأسرية وغيرها.
- ❖ إعادة الاختبار: (Test-Retest) أحد أساليب حساب ثبات الاختبار، وفيه يتم حساب الثبات عن طريق تكرار إعطاء الاختبار نفسه للمجموعة الطلابية نفسها بعد فترة زمنية محددة، ثم تحتسب معاملات الارتباط بين مجموعتي الدرجات للاختبارين.

❖ الانحراف المعياري: (Standard Deviation) يعد الانحراف المعياري أكثر مقاييس التشتت استخداماً، حيث يعتمد في حسابه على إيجاد انحراف كل درجة من درجات توزيع معين عن متوسط التوزيع.

❖ البطارية: مجموعة مختارة بعناية من الاختبارات المترابطة التي تقيس سمة أو قدرة واحدة أو أكثر يتم عرضها على مجتمع معين من الأفراد، وعادة ما يتم تقييم مستوى الأداء على أحد اختباراتها أو على عدد منها أو جميعها، ومن بطارية الاختبار المشهورة بطارية GRE، وبطارية الاستعدادات الفارقة DAT، وبطارية التحصيل الدراسي SAT.

❖ تاريخ الحالة: (Case History) تلك البيانات والمعلومات التراكمية المتعلقة بفرد من الأفراد، كالخلفية الأسرية، والتاريخ الشخصي والنمو الجسمي، والتاريخ الطبي، ونتائج الاختبارات، والسجلات القصصية للسلوك. ويستخدم تاريخ الحالة غالبا في اتخاذ القرارات المتعلقة بالواجبات المحددة للطالب وفى إعداد وتخطيط البرامج والخدمات اللازمة له

❖ التباين: (Variance) أحد مقاييس تشتت البيانات الكمية، وهو مجموع مربعات انحرافات القيم عن متوسطها مقسوماً على عدد القيم التي يتألف منها مجتمع البيانات، أو مقسوماً على عدد القيم التي تتألف من عينة البيانات مطروحاً منها واحد.

❖ تجريب أولي استطلاعي للاختبار: Pilot Testing التجريب المبدئي لعملية تقويم لا تزال في مرحلة المراجعة.تركز هذه العملية غالبا على ما يلي:

شمولية ووضوح إرشادات الاختبار، الشكل العام لمواد التقويم، مستوى برنامج تدريب المقيمين، والوقت المخصص للإختبار.

- ❖ التحصيل: وهو المستوى الذي تعلمه الفرد للقيام بالأداء على مهارة معينة. وعادة ما يرتبط التحصيل بمجمل المعلومات والمهارات والتمارين والأفكار التي اكتسبها خلال سنة أو مرحلة دراسية معينة.
- ❖ التحليل العاملي: (Factor Analysis) أسلوب رياضي يمثل عدداً كبيراً من العمليات والمعالجات الرياضية في تحليل الارتباطات بين المتغيرات (فقرات المقياس أو الاختبار)، ومن ثم تفسير هذه الارتباطات واختزالها في عدد أقل من المتغيرات تدعى عوامل. ويساعد التحليل العاملي في التوصل إلى أن السمة المقاسة أحادية البعد أو متعددة الأبعاد
- ❖ التشتت: (Variation) حالة تباعد البيانات عن بعضها بعضاً، أو تباعدها عن أحد مؤشرات النزعة المركزية.
- ❖ التشخيص: (Diagnosis) هو عملية منظمة تهدف إلى وصف وتحليل النتائج لتحديد العلة ومصدرها.
- ❖ التصنيف: (Classification) هو تلك العملية التي يمكن بها وضع الفرد في فئةٍ ما وفق خصائص أو مميزات مشتركة. بهدف تحديد نوع الخدمات التربوية والاجتماعية والطبية والمهنية اللازمة لكل فئة. وإعداد البرامج التي تناسب كل فئة على حدة بما يتناسب مع خصائصهم العقلية والجسمية والانفعالية والاجتماعية.

- تضخم الدرجات: (Grade Inflation) تعبير يشير إلى أن الدرجات التي تعطى للطلاب من قبل معلميهم لم تعد تعبر عن الأداء الحقيقي لهؤلاء الطلاب،وذلك بسبب إفراط المعلمين في منح تلك الدرجات.

- التقدير: يعرف التقدير بأنه عملية شاملة رسمية أو غير رسمية لجمع معلومات تتعلق بالأفراد وذلك بهدف الوصول الي فهم اعمق للافراد ، وقد لا تقتصر عملية التقدير علي استخدام الاختبارات ، او اتباع طرق القياس الرسمية بل قد تستخدم طرق أخري غير رسمية للحصول علي معلومات متنوعة حول الافراد مثل السجلات المدرسية او المقابلات الشخصية أو غيرها من المصادر الأخري، وتتطلب عملية التقدير جمع للمعلومات حول الشيئ الذي نقوم بتقديره من مصادر متنوعة ولا تتطلب الحكم علي جدوي أو قيمة هذه المعلومات (Wiersma&Jurs 1990) و(علي ماهر خطاب 2001)

- التقويم: هناك من اعتبر التقويم والتقييم بمعنى واحد، وهناك من فرق بينهما حيث اعتبر التقويم بمعنى تعديل، والتقييم بمعنى تثمين وحكم)- التقويم لغة التعديل والتحسين، تقول: قومت كذا أي عدلته وحسنته، وهي الحكم والتثمين تقول أقومها بكذا. إصطلاحاً: هو إعطاء حكم بناءَ على وصف كمي أو كيفي. أو هو إعطاء (إصدار) حكم على الظاهرة المراد قياسها في ضوء ما تحتوي من خصائص.

- ❖ التقييم: إصدار حكم على قيمة الأشياء أو الأفكار أو الجوانب أو الاستجابات لتقدير مدى كفاية هذه الأشياء ودّقتها وفاعليّتها، على أن يتمّ هذا الحكم في ضوء مستوى أو محكّ أو معيار معيّن، ويتمّ تقييم الشخص في ضوء المستويات المحدّدة من قبل مثل: ضعيف، مقبول، جيّد، جيّد جدّاً، ممتاز.
- ❖ ثبات الاختبار: يقصد بثبات الاختبار دقته واتساقه، وبمعنى أدق ان يعطي الاختبار نفس النتائج إذا ما تم استخدام أكثر من مرة تحت ظروف مماثلة.

ثبات الفقرة: (Item Reliability) هو الارتباط بين درجات الافراد على المفردة والدرجة الكلية على الاختبار ويتم حسابهان طريق مايلي :

أ- باستخدام معامل الارتباط الثنائى

ب- باستخدام طريقة الاحتمال المنوالي

ت- باستخدام طريقة إعادة الاختبارومعادلة معامل ارتباط فاي

- ❖ الخاصية: مفهوم ادراكي وكمي لا يتوفر في كل الأوقات ولكل الناس بالرغم من وجودها في حياتنا اليومية (الجاذبية، والأحلام، وزلات اللسان، والاستجابات الشرطية المختلفة...الخ) كما أنها تعبر عن صفة أو نوعية أو ميزة يفترض وجودها عند عامة الناس أو فئة من الاشياء مع افتراض أنها لا توجد بنفس الكمية (الطول، الوزن، الذاكرة، التفكير، الميل) وهي أيضا بسيطة ومركبة، ومنفردة ومتداخلة.

- ❖ الخطأ العشوائي: (Random Error) هو مجمل ما يؤثر على الدرجات الخام للمفحوصين من مصادر مختلفة يصعب ضبطها والحد من تأثيرها.
- ❖ خطأ القياس المعياري: (Standard Error of Measurment) مقياس للخطأ المتوقع في درجة الاختبار، أنه المقدار الذي في حدوده تزيد أو تنقص الدرجه المشاهدة عن الدرجة الحقيقية للطالب. وكلما قل خطأ القياس المعياري كلما زادت الدقة في درجات الاختبار. خطأ القياس المعياري هو عبارة عن الانحراف المعياري لمجموعة من الفروقات كل منها يساوي الفرق بين الدرجة الحقيقية والدرجة المتحصل عليها. الخطأ المعياري له علاقة مباشرة بثبات الاختبار، فكلما كانت درجة ثبات الاختبار عالية كلما قل خطأ القياس
- ❖ خطأ القياس: (Error of Measurement) مقدار الفرق بين الدرجة التي يحصل عليها المفحوص (الدرجة الخام أو الملاحظة) والدرجة المفترضة (الحقيقية). ويقع خطأ المقياس في نوعين: عشوائي ومنتظم.
- ❖ الخطأ المعياري في التقدير: (Standard Error of Estimate) هو الانحراف المعياري المعتمد على الفروق بين الدرجات الملاحظة والدرجات المتنبأة من خلال معرفة معامل الارتباط بين الدرجات المتنبأة والدرجات على المحك.
- ❖ الخطأ المعياري للقياس: (Standard Error of Measurement) هو مقدار التذبذب المتوقع للدرجة الخام حول الدرجة الحقيقية.

- ❖ الخطأ المنـتظم: (Systematic Error) مقـدار مـا يـؤثر عـلى درجـة المفحـوص مـن مصادر يمكن معرفتها وبالتالي ضبطها.
- ❖ الدرجـة التائيـة: (T - Scores) هـي درجـة معياريـة في توزيـع متوسـطه الحسـابي يساوي (50) وانحرافه المعياري يساوي (10). وتمتد قـيم الـدرجات التائيـة للتوزيـع الواحد بين (صفر) و (100).
- ❖ درجة الحقيقية: (True Score) هي درجة افتراضية خالية تماما من الخطـأ حققها طالب على اختبار. يمكن تصور الدرجة الحقيقية عـلى أنهـا متوسـط الدرجة التي يحصل عليها الطالب عندما يتقدم للإختبار نفسه عدد لا نهائي من المرات، شريطة الا يحدث بين هذه الاختبارات نضج للطالب يؤثر عـلى إجابته.
- ❖ درجة الخام: (Raw Score) عدد الأسئلة التي حقق الطالب فيها إجابـات صـحيحة، ورغم أهمية الدرجة الخام إلا إنها يجب ألا تستخدم لعمل مقارنات عـلى اختبـارات مختلفة. الدرجة الخام هي درجة الطالـب عـلى الاختبـار قبـل أن يـتم معالجتهـا أو تعديلها إحصائيا.
- ❖ الدرجة الزائية: (Z - Score) هي درجة معياريـة في توزيـع متوسـطه الحسـابي يساوي (صفراً) وانحرافه المعياري يساوي (1). وبذلك يمكن مقارنة الدرجة مـع غيرها من الدرجات التي تنتمي لنفس التوزيع، وذلك بحساب الدرجة المعيارية الزائية لكل منها، ومن ثم يكون التفضيل في ضوء مقدار قيمة الدرجة المعيارية (z).

❖ درجة المتوسط: (Average Score) إحدى الطرق العديدة للتعبير عن مستوى أداء مجموعة من الطلاب، يتم الحصول على الدرجة المتوسطة بقسمة مجموع درجات الطلاب على عددهم، يطلق على هذه الدرجة أحيانا الوسط الحسابي (Arithmetic mean). تتأثر الدرجة المتوسطة بالدرجات المتطرفة (العالية أو المنخفضة).

❖ الدرجة المعيارية: ترتبط فكرة الدرجات المعيارية بالتوزيع الطبيعي المعياري (القياسي) والذي يتميز بأن انحرافه المعياري واحد صحيح ومتوسطه صفر، وبالتالي يمكننا أن نحول أي مجموعة من الدرجات ذات توزيع طبيعي إلى توزيع طبيعي معياري، باستخدام الدرجات المعيارية والتي يتم حسابها بقسمة انحراف الدرجة عن متوسطها على الانحراف المعياري للدرجات. وهذه الفكرة تتيح لنا مقارنة أي مجموعة من الدرجات بمجموعة أخرى مختلفة عنها، ويصبح لمجموعتي الدرجات نفس المتوسط ونفس الانحراف المعياري بعد تحويل مجموعتي الدرجات إلى درجات معيارية. وجدير بالذكر أن وحدة الدرجات المعيارية هي الانحراف المعياري، ويمكن التخلص من الإشارات السالبة في الدرجات المعيارية باستخدام الدرجات المعيارية المعدلة، وذلك بضربها في 10 ليصبح انحرافها المعياري 10، وإضافة قيمة ثابتة لكل الدرجات لتكون هي المتوسط الجديد.

❖ درجة المئينية: (Percentile Rank) النسبة المئوية من توزيع الدرجات التي تقع تحت درجة معينة، وهي درجة تعبر عن مستوى أداء الطالب على

الاختبار بالنسبة لزملائه. الدرجة المئينية لطالب تعني النسبة المئوية من الطلاب الذين درجاتهم تساوي أو تقل عن درجته. فمثلا، إذا وجد أن 70% من الدرجات تقل عن الدرجة الخام 56 فإن الدرجة 56 تقع عند السبعين المئينية (th Percentile70). الطالب الذي حصل على درجة مئنية مقدارها 63 percentile63rd – مثلا- يكون قد حقق أداء مساويا أو أفضل من 63% من الطلاب الذين تم اختبارهم. الطلاب الموهوبون يحققون عادة رتب مئينية في حدود 90 أو 95، أي أن درجاتهم تتفوق على 90% أو 95% من زملائهم.

- ❖ درجة تمييز الفقرة: (Discrimination Index) درجة تمييز الفقرة تدل على قدرة الفقرة أو السؤال على التمييز بين الطلاب ذوي التحصيل العالي والطلاب ذوي التحصيل المتدني، وتعتبر درجة تمييز الفقرة من أهم المؤشرات التي تصف جودة بنود الاختبار، وذلك نظرا لأن أهم وظائف الاختبار هي أن يميز بين الطلاب الضعاف أكاديميا والطلاب المتفوقين. وتتراوح درجات التمييز بين +1 و -1، وبثبات كل شيء آخر فإن ارتفاع درجة التمييز للفقرة يدل على جودتها.يعتبر3 فأكثر تمييز جيد، ويعتبر2 – 3 تمييز مقبول، وأقل من 2 لا يميز
- ❖ درجة: (Score) أي رقم محدد ينتج عن تقييم لقدرات الفرد وإمكاناته. يعبر هذا الرقم غالبا عن درجة إمتلاك المتعلم للمهارة التي يتم قياسها على سلم الدرجات.

- دليـل الاختبـار: (Test Manual) منشـور يعـده مطـور الاختبـار والنـاشر ليعطـي معلومات عن أسـلوب إدارة الاختبـار وتصـحيحه وتفسـير مـدلول الـدرجات، يقـدم المنشور أيضا بيانات علمية عـن خـواص الاختبـار والإجـراءات المسـتخدمة في بنـاءه وتطويره.
- دليل التصحيح: (Scoring Guide) مجموعة الخطوط العامة التـي تصـف مـا يـتم تقييمه وتقرر أسلوب التقيـيم ومـا يجـب أن يلتـزم بـه المصحح عنـد تقيـيم عمـل الطالب.
- سلم ليكارت: هذا السلم شائع في البحوث النفسية والتربوية، يتدرج هذا السلم في مستوى الصفة من (1-5)، مثـل: أوافـق بشـدة، أوافـق، مـتردد، لا أوافـق، لا أوافق بشدة)، ثم تترجم الألفاظ إلى أرقام ويتعامل معها إحصائياً.ويمكـن عمـل سلم بثلاث خيارات كما يمكن عمل سلم بسبع خيارات واكثر. ولكن كلـما زادت الخيارات كلما كانت النتائج أدق، ولكن الجهـد الإحصـائي شـاق، لـذا كثـير مـن الباحثين يفضلون الخيارات الثلاث رغم أنها أقل دقـة، لكـن لأنهـا لا تحتـاج إلى جهد أكبر.
- السمة: (Trait) مفهوم يستخدم في وصف سلوك الأفراد، وهو عبـارة عـن سـلوك أو مجموعة من السلوكات التي يميل الفر د إلى إظهارها. فالسمة ليست صـفة مفـردة، وليست شيئاً ملموساً بل هي مفهوم مجرد.

- ❖ صدق الاختبار: يقصد بصدق الاختبار مدى قدرته على قياس المجال الذي وضع من أجله أو بمعنى أكثر تحديدا مدى صلاحية درجاته للقيام بتفسيرات مرتبطة بالمجال المقاس.
- ❖ عينة الدراسة: مجموعة جزئية من مجتمع الدراسة يتم اختيارها وفق شروط معينة لتمثل المجتمع الأصلي.
- ❖ الفقرة: هي السؤال أو العبارة أو التمرين أو المهمة أو المسالة التي يتضمنه الاختبار أو المقياس أو أداة التقييم،وتصاغ الفقرات بأكثر من نمط، فمنها الفقرات الانتقائية ويطلب فيها من المفحوص أن يختار الإجابة الصحيحة من بين عدد من البدائل)، والفقرات الصياغية (ويطلب فيه من المفحوص أن يصيغ الإجابة الصحيحة على السؤال).
- ❖ القائمة: هي دليل أو قائمة من المفردات أو العبارات تهدف إلى تقييم مدى وجود اتجاهات أو اهتمامات أو سلوكيات محددة لدى الأفراد، وغالبا ما تعرف كأدوات ورقة وقلم و تستخدم في تقويم الشخصية أو الاتجاهات أو الميول، وتعرف أحيانا بأدوات التقرير الذاتي، ونادرا ما تستخدم كمقاييس تحصيل أو مهارات أو تحصيل أو قدرات أو تحصيل.
- ❖ القدرة: هي كفاية الفرد في مجال محدد، وهي خاصية من خصائص الفرد. وغالبا ما يختلط الامر بين القدرة والاستعداد، فبينما ترتبط القدرة بالحالة القائمة للشخص يرتبط الاستعداد بالواقع المستقبلي أو التنبؤي له.

❖ القياس: القياس بمفهومه الواسع:هو التعبير عن الأشياء بأعداد حسب قواعد محددة (القياس هنا هو عدد الوحدات التي تمثل الخاصية) ومفهومه في التربية: هو عملية تعتمد على جمع معلومات من أجل تقدير الأشياء تقديراً كمياً في ضوء وحدة قياس معينة. أو هو العملية التي نحدد بواسطتها كمية ما يوجد بالشيء من الخاصية أو السمة التي نقيسها.

❖ الكفاية: مستوى من الأداء المرضي في القيام بمقتضيات العمل، والتي تتطلبها مهنة معينة بالذات.

❖ المتغيرات: Variables يقصد بالمتغير " أي خاصية يمكن قياسها وتتباين قيمها من فرد إلى آخر أو من مجموعة إلى أخرى".

❖ المتوسط الحسابي: (Arithmetic Mean) نقطة توزيع من القيم الكمية، يعرف أحياناً بالمعدل (Average). من ناحية أخرى، يساوي مجموع انحرافات كافة القيم عن المتوسط الحسابي صفراً. ويتم حسابه كمياً بإيجاد حاصل قسمة مجموع القيم على عددها.

❖ مجتمع الدراسة: كامل الأفراد أو الأحداث أو المشاهدات موضوع البحث أو الدراسة التي يسعى الباحث إلى أن يعمم عليها النتائج ذات العلاقة بالمشكلة المدروسة.

❖ المحك: (Criterion) مجموعة (مجال) من المعارف والمهارات المحددة تحديداً جيداً بحيث يمكن نتيجة لمقارنة أداء الفرد في الاختبار بهذا المجال أن نعرف ما يستطيع الفرد أن يؤديه وما لا يستطيع.

- محكات الأداء: (Performance Criteria) هي المستويات التي يتم تقييم أداء الطالب في ضوئها. وتساعد هذه المعايير المقيمين على تحقيق الموضوعية من ناحية، وتزويد المتعلم بمعلومات مهمة عن التوقعات التي تقود إلى الهدف أو الغاية المنشود تحقيقها من ناحية أخرى.
- المدى: (Range) أحد مقاييس تشتت البيانات الكمية، وهو الفرق بين أكبر قيمة وأصغر قيمة في التوزيع (لذلك فهو كمية موجبة دائماً).
- مستويات القياس (موازين القياس): عبارة عن تصنيف الخصائص والمتغيرات التربوية والنفسية، وفقا لأسلوب التعامل مع القيم الرقمية إلى أربعة مستويات قياسية، هي: القياس الاسمي، والقياس الرتبي، والقياس الفئوي، والقياس النسبي.
- معامل الصعوبة: معامل الصعوبة يعد أحد أهم المؤشرات التي ينظر فيها التربويون عندما يجرون تحليلا لبنود الاختبار، يفيد معامل الصعوبة في إيضاح مدى سهولة أو صعوبة سؤال ما في الاختبار، وهو عبارة عن النسبة المئوية من الطلاب الذين أجابوا عن السؤال إجابة صحيحة ويحسب بتطبيق المعادلة التالية: عدد الطلاب الذين أجابو على السؤال إجابة صحيحة/ مجموع الطلاب × 100.

معامل التمييز للفقرة: (Item Discrimination Index) قدرة الفقرة على التمييز بين المفحوصين من فئة ذوي الأداء المنخفض وذوي الأداء المرتفع في إجابتهم على الفقرة.

معامل التمييز = عدد الإجابات الصحيحة عن الفقرة في المجموعة العليا _ عدد الإجابات الصحيحة عن الفقرة في المجموعة الدنيا ÷ عدد أفراد أحد المجموعتين .

وبهذا يكون معامل التمييز = $\frac{\text{س - ص}}{\text{ن}}$

حيث س : عدد طلاب الفئة العليا في التحصيل الذين أجابوا على السؤال إجابة صحيحة.

ص: عدد طلاب الفئة الدنيا في التحصيل ممن أجابوا على السؤال إجابة صحيحة.

ن : عدد افراد إحدى المجموعتين.

- معايير التصحيح: (Scoring Criteria) مجموعة من المحكات التي يضعها المصحح مسبقا ليحكم في ضوئها على جودة إجابة الطالب عن أي سؤال ورد في الاختبار.
- المعايير: المعيار مفرد المعايير، وكثيرا ما يستخدم مصطلح المعايير لتعبير عن مجموعة معيارية في القياس التربوي،كما يستخدم مرادفا للمعدل الذي هو الدرجة الوسيط لمجموعة معينة من الناس. وهذه المجموعة المعينة تسمى بالمجموعة المعيارية أو المرجعية.
- المفهوم الإجرائي: وصف الشيء أو الموضوع أو الخاصية وصفا إجرائيا من خلال بيان المواقف السلوكية التي يظهر عليها.

- المقابلة: (Interview) هي أسلوب علمي منظم لجمع البيانات والمعلومات عن السلوك أو الظاهرة موضوع القياس.
- المنوال: (Mode) القيمة ذات التكرار الأكبر في توزيع البيانات، وقد يحوي بعض التوزيعات أكثر من منوال. عند ذلك يدعى التوزيع بثنائي أو ثلاثي أو حتى متعدد المنوالات. وقد يكون المنوال مقداراً كمياً أو صفة نوعية.
- المهارة: وهو المستوى الذي يبلغه الفرد في أداء عمليات حركية جسمية معقدة، يتوفر فيها عناصر السرعة والدقة والتكيف مع الظروف المتغيرة.
- موضوعية الاختبار: ويقصد بموضوعية الاختبار، عدم تأثر المصحح بالعوامل الذاتية عند تصميمه لأوراق الإجابة.
- المئينات: (Percentiles) هي القيمة التي تقسم عينة من البيانات الكمية إلى مائة قسم متساوية، كل واحد يحوي عدداً متساوياً (بدرجة كبيرة) من البيانات.
- هدف سلوكي: (Behavioral Objective) عبارة تعليمية أو تدريسية تشتمل على السلوك المتوقع والشروط التي يتم فيها ذلك السلوك ومستوى التمكن أو الأداء المطلوب.
- الوسيط: (Median):نقطة في توزيع القيم الكمية، وهي النقطة التي يكون (50 %) من القيم مرتبة تصاعدياً أو تنازلياً و (50 %) بعدها. ويعرف الوسيط أيضاً بأنه المئين (50).

قائمة المراجع:

ابراهيم، المحاسنة (2013). القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة. دار جرير للنشر والتوزيع، عمان.

أبو لبدة، سبع محمد (1982). "مبادئ القياس النفسي والتقييم التربوي"، ط2، الجامعة الأردنية، عمان.

أبو لبده، سبع (1999). مبادئ القياس النفسي والتقييم التربوي، الطبعة الأولى الجامعة الأردنية، عمان.

أبو هاشم حسن، السيد (2006). الخصائص السيكومترية لأدوات القياس فى البحوث النفسية والتربوية باستخدام SPSS ، كلية التربية، جامعة الملك سعود.

أحمد، محمد الطيب(1999). التقويم والقياس والنفسي والتربوي، المكتب الجامعي الحديث الإسكندرية، مصر.

احمد، محمد عبدالرحمن (2011). تصميم الاختبارات - اسس نظرية وتطبيقات عملية. دار اسامة للنشر والتوزيع، عمان.

تيسير، مفلح كوافحة (2003). القياس والتقييم واساليب القياس والتشخيص في التربية الخاصة. دار المسيرة للنشر والتوزيع والطباعة، عمان.

جامعة القدس المفتوحة، (1994). "القياس والتقويم"، برنامج التربية، عمان.

الجبالي، حسنى (2008). الفروق الفردية فى القدرات العقلية والقياس والتقويم النفسي والتربوي، القاهرة، مكتبة الانجلو المصرية.

جودت، عزت عبدالهادي (2002). مبادئ القياس والتقويم في التربية. الدار العلمية الدولية، عمان.

الحازمي، عدنان (2007). الإعاقة العقلية : دليل المعلمين وأولياء الأمور. دار الفكر، عمان – الأردن.

الحسيني، طه الشربيني (2001). القياسات والمقاييس الالكترونية. دار الراتب الجامعية، بيروت.

ذوقان، عبيدات؛ كايد، عبد الحق؛ و عبد الرحمن، عدس(2004). البحث العلمي مفهومه وأدواته وأساليبه.

راشد، الدوسري (2004). القياس والتقويم التربوي الحديث مبادئ وتطبيقات وقضايا معاصرة. دار الفكر للنشر والتوزيع، عمان, الأردن.

روبرت، ثورندايك (1986). القياس والتقويم في علم النفس والتربية. مركز الكتب الاردني، عمان.

الروسان، فاروق (2006). أساليب القياس والتشخيص في التربية الخاصة. ط 2، دار الفكر للطباعة والنشر والتوزيع، عمان، الأردن.

زكريا، محمد الظاهر (2002). مبادئ القياس والتقويم في التربية، دار العلمية الدولية للنشر والتوزيع، عمان، الأردن.

زيتون، حسن (2007). أصول التقويم والقياس التربوي : المفهوم والتطبيقات، الدار الصولتية للتربية، الرياض، المملكة العربية السعودية.

زيد، الهويدي (2004). أساسيات القياس والتقويم التربوي، دار الكتاب لجامعي العين، الإمارات العربية المتحدة.

سامي، محمد ملحم (2002). القياس والتقويم في التربية وعلم النفس. ط2، دار المسيرة للنشر والتوزيع الأردن.

سليمان، عبدالرحمن (2004). معجم التخلف العقلي. مكتبة زهراء الشرق، القاهرة – مصر.

سمارة، نواف؛ والعديلي، عبدالسلام (2008). مفاهيم ومصطلحات في العلوم التربوية. دار المسيرة، عمان – الأردن.

سوسن، شاكر مجيد (2010). الاختبارات النفسية – نماذج. دار صفاء للنشر والتوزيع، عمان.

شين، لوبيز سنايدر (2013). القياس في علم النفس الايجابي، نماذج ومقاييس. المركز القومي للترجمة، القاهر.

صالح، حسن الداهري (2011). اساسيات القياس النفسي في الارشاد والصحة النفسية. دار الحامد للنشر والتوزيع، عمان.

صفوت، فرج (1980). القياس النفسي. الطبعة الأولى، دار الفكر العربي، القاهرة.

صلاح الدين، علام (1999). القياس والتقويم التربوي والنفسي أساسيته وتطبيقاته وتوجهاته المعاصرة، دار الفكر العربي القاهرة، مصر.

صلاح الدين، علام (2000): "القياس والتقويم التربوي والنفسي، أسسه وتطبيقاته، وتوجهاته المعاصرة"، دار الفكر العربي" القاهرة.

صلاح الدين، علام (2006). الاختبارات والمقاييس التربوية والنفسية.. دار الفكر للنشر والتوزيع، عمان, الأردن.

صلاح الدين، علام (2007). القياس والتقويم التربوي في العملية التدريسية. دار المسيرة للنشر والتوزيع والطباعة، عمان.

صلاح الدين، علام (2013). اتقان القياس النفسي الحديث - النظريات والطرق. دار الفكر للنشر والتوزيع، عمان, الأردن.

الظاهر، قحطان (2004). مصطلحات ونصوص إنجليزية في التربية الخاصة. دار اليازوري، عمان – الأردن.

عبده، أشرف (2006). الاتجاهات الحديثة في التقويم التربوي والقياس السيكولوجي، الدار الصولتية للتربية، الرياض – المملكة العربية السعودية.

عريفج، سامي؛ ومصلح، خالد ؛ وحواشين، مفيد (1982). مناهج البحث العلمي وأساليبه.الاردن، ط2، دار مجدلاوي للنشر والتوزيع.

عسكر،علي (1998). مقدمة في البحث العلمي. الكويت، ط3 مكتبة الفلاح.

علي، مهدي كاظم (2001). القياس والتقويم في التعلم والتعليم. دار الكندي للنشر والتوزيع والطباعة، اربد الاردن.

فاروق، الروسان (1999). اساليب القياس والتشخيص في التربية الخاصة، دار الفكر للنشر والتوزيع، عمان.

فاروق، السيد عثمان (1995). الاحصاء التربوي والقياس النفسي. دار المعارف للنشر والتوزيع، القاهرة.

فخري، رشيد خضر (2003). الختبارات والمقاييس في التربية وعلم النفس، الطبعة الأولى، دار القلم، دبي، الإمارات العربية المتحدة.

فرج، صفوت (2012). القياس النفسي , (ط7) , مكتبة الانجلو المصرية , مصر.

كمال الدين، محمد هاشم ؛ وحسن جعفر الخليفة (2011). التقويم التربوي، مفهومه، أساليبه، مجالاته، توجهاته الحديثة، ط 3، الرياض : مكتبة الرشد.

الكيلاني، عبدالله ؛ والروسان، فاروق (2006). التقويم في التربية الخاصة. دار المسيرة، عمان – الأردن.

ماهر، إسماعيل صبري؛ ,ومحب الرافعي (2008):التقويم التربوي أسسه واجراءاته ,الرياض : مكتبة الرشد.

محمد، جاسم العبيدي (2011). القياس النفسي والاختبارات. الاقلام الهادفة للنشر والتوزيع عمان.

محمد، حسن غانم (2007). القياس النفسي للشخصية. المكتبة المصرية، القاهرة.

محمد، صالح الامام (2011). القياس في التربية الخاصة - رؤية تطبيقية. الاقلام الهادفة للنشر والتوزيع، عمان.

محمد، عبدالهادي حسين (2003). قياس وتقييم قدرات الذكاء المتعددة. دار الفكر للنشر والتوزيع، عمان, الأردن.

محمود، عوض(1999). القياس النفسي بين النظرية و التطبيق، دار المعرفة الجامعية للنشر و التوزيع، عمان.

معزوز، علاونة (2014). معجم مصطلحات في القياس والتقويم النفسي والتربـوي (عـربي - انجليزي انجليزي - عربي). دار الشيماء للنشر والتوزيع، رام الـلـه، فلسطين.

مقدم، عبد الحفيظ (1997).الإحصاء و القياس النفسي و التربوي مع نماذج مـن المقـاييس، ديوان المطبوعات الجامعية.الجزائر.

النبهـان، مـوسى (2004). أساسـيات القيـاس في العلـوم السـلوكية. دار الشروق، عـمان – الأردن.

يعقوب، حسين نشوان (2011). القياس والتقويم وتطبيقاته في الـتعلم عـن بعـد والـتعلم الجامعي المفتوح، دار الفرقان، عمان.

يوسف، كماش (2013). القياس والاختبار والتقويم في المجال التربوي والرياضي. دار دجلـة ناشرون وموزعون، عمان.

خطاب، علي ماهر (2001) .القياس والتقويم في العلوم النفسية والتربوية والاجتماعية ط2 مكتبة الانجلو المصرية، الرابط:

http://www.socialresearchmethods.net/kb/measlevl.htm

المراجع الاجنبية:

Anastasi A (1982). Psychological Testing (5th ed.). New York: Macmillan.

Anderson, K.F.,(1994) .Constant Current Impedance Measuring System That Is Immune to the Effects of Parasitic Impedances, U.S. Patent No. 5,731,469, December.

Association for Assessment in Counseling and Education. (2001). Responsibilities of users of standardized tests (RUST) (3rd ed.). Online: http://aac.ncat.edu/documents/rust.html

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues.Educational Psychologist, 50(1), 84-94.

Bangertdrowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1983). Effects of coaching performance on achievment - test performance. Review of Educational Research, 53(4), 571-585.

Barney, G.C.,(1985). Intelligent Instrumentation, Microprocessor Applications in Measurement and Control, Prentice-Hall, Englewood Cliffs, NJ,.

Biesta, G. (2015). Education, Measurement and the Professions: Reclaiming a space for democratic professionality in education. Educational Philosophy and Theory, (ahead-of-print), 1-16.

Bond, T. G., & Fox, C. M. (2013). Applying the Rasch model: Fundamental measurement in the human sciences. Psychology Press.

Braun, H., Carlson, S., & Bejar, I. I. (1989). Psychometric foundations of testing based on patient management problems (RM-89-2).Princeton, NJ:

Bryon, M., Clayden, J., & Tyreman, C. J. (2010). How to master the UKCAT : Over 750 practice questions for the United Kingdom clinical aptitude test (2nd ed. ed.). London: Kogan Page.

Butterworth, J., & Thwaites, G. (2010). Preparing for the BMAT: The official guide to the Biomedical Admissions Test Portsmouth, NH: Heinman.

Canadian Psychological Association. (2004). Ethical use and reporting of psychological assessment results for student placement. CPA policy statement 2004-02. Ottawa: CPA.Online: http://www.cpa.ca/aboutcpa/policystatements/#ethical

Carroll, J. B. (1980). Measurement of abilities constructs. In Construct Validity in Psychological Measurement. Proceedings of a colloquium on theory and application in education and employment. Princeton, NJ: Educational Testing Service.

Christensen, M. (2001). Diagnostic criteria in clinical settings: DSM-IV and cultural competence. American Indian and Alaska Native Mental Health Research.

Cook, M. (2006). Personnel selection: Adding value through people. Chichester: Wiley.

Costa PT, McCrae R (1992) Revised NEO Personality Inventory (NEO-PI-R) (NEO-FFI) Professional Manual. Odessa, FL: Psychological Assessment Resources.

Dockrell, J. E., & Marshall, C. R. (2015). Measurement Issues: Assessing language skills in young children. Child and Adolescent Mental Health, 20(2), 116-125

Donlon, T. F. (Ed.). (1984). The College Board technical handbook for the Scholastic Aptitude Test and Achievement Tests. New York: College Entrance Examination Board.

Dunn-Rankin, P., Knezek, G. A., Wallace, S. R., & Zhang, S. (2014). Scaling methods. Psychology Press.

Ebel, R. (1979). Instructor's Manual: Essentials of Education Measurement. Prentice-Hall.

Emery, J. L., & Bell, J. F. (2009). The predictive validity of the BioMedical Admissions Test for pre-clinical examination performance. Medical Education, 43(6), 557-564.

Emery, J. L., Bell, J. F., & Rodeiro, C. L. V. (2011). The BioMedical Admissions Test for medical student selection: Issues of fairness and bias. Medical Teacher, 33(1), 62-71.

Eysenck, H. J., & Eysenck, S. B. (2013). Personality Structure and Measurement (Psychology Revivals). Routledge.

Flanagan, D. P., & Ortiz, S. O. (2001). Essentials of cross-battery assessment. New York, NY: John Wiley & Sons, Inc.

Friedman, A. F., Bolinskey, P. K., Levak, R. W., & Nichols, D. S. (2014).Psychological Assessment with the MMPI-2/MMPI-2-RF. Routledge.

Garb HN (1998) Studying the Clinician: judgement research and psychological assessment. Washington, DC: American Psychological Association.

Gardner, H. (1983). Frames of Mind: The Theory of Mulltiple Intelligences. New York: Basic Books.

Gardner, H. (1992). Multiple intelligencies: The theory in practice. New York: Basic Books.

Green, M., & Hawley, N. (2009). Succeeding in the Bio Medical Admissions Test 2009 (BMAT) (Entry to Medical School). London: Apply2 Ltd.

Gregory, R. J. (1994). Aptitude tests. In R. J. Sternberg (Ed.), Encyclopedia of human intelligence. New York: Macmillan.

Gregory, R. J. (2010). Psychological testing : history, principles, and applications (6th ed.). Boston, Mass.: Allyn and Bacon.

Guildford, J. P. (1985). The structure-of-intellect model. In B. B. Wolman (Ed.), Handbook of intelligence: Theories, measurements and applications. New York: Wiley.

H.M.Government. (April 2011). Opening Doors, Breaking Barriers: A Strategy for Social Mobility.

Hathaway SR, McKiney JC (1983) The Minnesota Multiphasic Personality Inventory (Manual). New York: Psychological Corporation.

Hays, P. A. (2001). Addressing cultural complexities in practice. A framework for clinicians and counsellors. Washington, DC: American Psychological Association.

Herrnstein, R., & Murray, C. (1994). The bell curve: Intelligence and class structure in American Life. New York: Free Press.

Higgins, L. T., & Sun, C. H. (2002). The development of psychological testing in China. International Journal of Psychology, 37(4), 246-254. 109

Huffcutt, A. I., & Roth, P. L. (1998). Racial group differences in employment interview evaluations. Journal of Applied Psychology, 83(2), 179-189.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.

Irvine, S. H., Dunn, P. L., & Anderson, J. D. (1989). Towards a theory of algorithm-determined cognitive test construction (Report).Devon, UK: Polytechnic South West.

Jensen, A. R. (1980). Bias in mental testing. New York: Free Press.

Joint Committee on Standards for Educational and Psychological Testing of the AERA, APA, and NCME. (1999). Standards for educational and psychological testing. Washington DC: American Educational Research Association.

Joint Committee on Standards for Educational Evaluation. (1994). The program evaluation standards: how to assess evaluations of educational programs (2nd ed.). Thousand Oaks, CA: Sage Publications. Educ Library – Stacks: LB3051 .J57 1994

Joint Committee on Standards for Educational Evaluation. (2003). The student evaluation standards: how to improve evaluations of students. Thousand Oaks, CA: Sage Publications. Educ Library – Stacks: LB3051 .J575 2003.

Joint Committee on Testing Practices. (2004). Code of fair testing practices in education. Washington, DC: APA, JCTP. Online: http://www.apa.org/science/programs/testing/fair-testing.pdf

Julian, E. R. (2005). Validity of the Medical College Admission Test for predicting medical school performance. Academic Medicine, 80(10), 910-917.

Kaplan, R., & Saccuzzo, D. (2012). Psychological testing: Principles, applications, and issues. Cengage Learning.

Khenissi, M. A., Essalmi, F., Jemni, M., & Chang, T. W. (2015). Measuring learners' working memory capacity from their interactions within educational game. In Emerging Issues in Smart Learning (pp. 233-237). Springer Berlin Heidelberg.

Kiresuk, T. J., Smith, A., & Cardillo, J. E. (2014). Goal attainment scaling: Applications, theory, and measurement. Psychology Press.

Kirkup, C., Wheater, R., Morrison, J., Durbin, B., & Pomati, M. (2010). Use of an aptitude test in university entrance: a validity study: Final Report: National Foundation for Educational Research.

Kleinman, A. (1996). How is culture important for DSM-IV? In J.E. Mezzich, A. Kleinman, H. Fabrega, & D.L. Parron, (Eds.), Culture and psychiatric diagnosis: A DSM-IV perspective (pp. 15-25). Washington, DC: American Psychiatric Press.

Kline P (2000) The Handbook of Psychometric Testing (2nd ed.). London, UK: Routledge.

Kulik, J. A., Bangertdrowns, R. L., & Kulik, C. L. C. (1984). Effectiveness of coaching for aptitude-tests. Psychological Bulletin, 95(2), 179-188.

Loh, C. S., Sheng, Y., & Ifenthaler, D. (Eds.). (2015). Serious Games Analytics: Methodologies for Performance Measurement, Assessment, and Improvement. Springer.

McCullough, C. S., Walker, J. L., & Diessner, R. (1985). The use of Wechsler scales in the assessment of Native Americans in the Columbia River Basin. Psychology in the Schools..

McManus, I. C., Ferguson, E., Wakeford, R., Powis, D., & James, D. (2011). Predictive validity of the Biomedical Admissions Test: An evaluation and case study. Medical Teacher, 33(1), 53-57. 110

Messick, S. (1981). Constructs and their vicissitudes in educational and psychological measurement. Psychological Bulletin, 89,575–588.

Michell J (1997) Quantitative science and the definition of measurement in psychology. British Journal of Psychology, 88, 355–383.

Michell, J. (2014). An introduction to the logic of psychological measurement. Psychology Press.

Mislevy, R. J., & Sheehan, K. M. (1990). How to equate tests with little or no data. Presented at the Psychometric Society Meeting. Princeton, NJ:.

MMY (2003) The Mental Measurements Yearbook (15th ed.). [buros.unl.edu/buros/catalog.html]

Morris, S. B., & Lobsenz, R. E. (2000). Significance tests and confidence intervals for the adverse impact ratio. Personnel Psychology, 53(1), 89-111.

Naglieri, J. A., LeBuffe, P. A., & Ross, K. M. (2013). Measuring resilience in children: From theory to practice. In Handbook of resilience in children (pp. 241-259). Springer US.

Pedhazur, E. J., & Schmelkin, L. P. (2013). Measurement, design, and analysis: An integrated approach. Psychology Press.

Picard, O. (2009). Get into medical school : 600 UKCAT practice questions : includes full mock exam, comprehensive tips, techniques and explanations. London: ISC Medical.

Ployhart, R. E., Schneider, B., & Schmitt, N. (2006). Staffing organizations: Contemporary practice and theory (3rd ed.). Mahwah, N.J.: Lawrence Erlbaum.

Polikoff, M. S., Zhou, N., & Campbell, S. E. (2015). Methodological choices in the content analysis of textbooks for measuring alignment with standards.Educational Measurement: Issues and Practice.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance - not much more than g. Journal of Applied Psychology, 79(4), 518-524.

Reeve, C. L., & Hakel, M. D. (2002). Asking the right questions about g. Human Performance, 15(1-2), 47-74.

Rogler, L. H. (1993). Culture in psychiatric diagnosis: An issue of scientific accuracy. Psychiatry.

Rust, J., & Golombok, S. (2014). Modern psychometrics: The science of psychological assessment. Routledge.

Salgado, J. F., & Anderson, N. (2002). Cognitive and GMA testing in the European community: Issues and evidence. Human Performance, 15(1-2), 75-96.

Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. Human Performance, 15(1-2), 187-210.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin, 124(2), 262-274.

Schultz, M. S., & Zedeck, S. (2008). Idenfication, development, and validation of predictors for successful lawyering. Retrieved from http://ssrn.com/abstract=1442118 website: 111

Schwartz, S. (2013). Measuring reading competence: A theoretical-prescriptive approach. Springer Science & Business Media.

Simien, E. (1986). Law School Admission Test as a Barrier to Almost Twenty Years of Affirmative Action. Thurgood Marshall Law Review, 12, 359-393.

Sternberg, R. J. (1986). Intelligence applied: Understanding and increasing you intellectual skills. San Diego, CA: Harcourt Brace Javanovich.

Stilwell, L. A., Dalessandro, S. P., & Reese, L. M. (2009). Predictive Validity of the LSAT: A National Summary of the 2007 and 2008 LSAT Correlation Studies: Law School Admission Council.

Thorndike, R. L. (1986). The role of general ability in prediction. Journal of Vocational Behavior, 29(3), 332-339.

Tyreman, C. J. (2009). How to master the BMAT: Unbeatable preparation for success in the Biomedical Admissions Test. London: Kogan Page.

Urbina, S. (2014). Essentials of psychological testing. John Wiley & Sons.

Van de Vijver, F. (2000). The nature of bias. In R. H. Dana (Ed.), Handbook of cross-cultural and multicultural personality assessment. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Weekley, J. A., & Ployhart, R. E. (Eds.). (2013). Situational judgment tests: Theory, measurement, and application. Psychology Press.

Whorton, J. W., & Morgan, R. L. (1990). Comparison of the test of nonverbal intelligence and Wechsler Intelligence Scale for Children-Revised in rural Native American and White children. Perceptual and Motor Skills, 70, 12-14.

Wright, S. R., & Bradley, P. M. (2010). Has the UK Clinical Aptitude Test improved medical student selection? Medical Education, 44(11), 1069-1076.

Wiersma , w. , Jurs,s.(1990).Educational measurments and testing .Boston :Allyn and Bacon . A division of Simon and Schuster Inc.

Yates, J., & James, D. (2010). The value of the UK Clinical Aptitude Test in predicting pre-clinical performance: a prospective cohort study at Nottingham Medical School. BMC Medical Education, 10.

Zimprich, D. (2015). Attitudes toward statistics among Swiss psychology students. Swiss Journal of Psychology.

بناء وتصميم الاختبارات

والمقاييس النفسية والتربوية

دار أمجد للنشر والتوزيع

جوال :
٠٠٩٦٢٧٩٦٩١٤٦٣٢
٠٠٩٦٢٧٩٩٢٩١٧٠٢
٠٠٩٦٢٧٩٦٨٠٣٦٧٠

هاتف :٤٦٥٢٢٧٢ ٠٠٩٦٢٦
فاكس:٤٦٥٣٣٧٢ ٠٠٩٦٢٦

dar.almajd@hotmail.com
dar.amjad2014dp@yahoo.com

عمان - الأردن - وسط البلد- مجمع الفحيص - الطابق الثالث



